Shaorong Sun

Five Basic Institution Structures and Institutional Economics

Science Press
Beijing

Springer

Five Basic Institution Structures and Institutional Economics

Shaorong Sun

Five Basic Institution Structures and Institutional Economics

Science Press
Beijing

Springer

Shaorong Sun
Management School
University of Shanghai for Science
and Technology
Yangpu District, Shanghai
China

ISBN 978-981-10-0343-1 ISBN 978-981-10-0344-8 (eBook)
DOI 10.1007/978-981-10-0344-8

Jointly published with Science Press, Beijing

Library of Congress Control Number: 2015959576

Printed on acid-free paper

This Springer imprint is published by SpringerNature
The registered company is Springer Science+Business Media Singapore Pte Ltd.

Preface

Two Decades of Research on Innovation

In 1995, with the publication of a simple article by the author in *Theories and Practice of Systems Engineering*, the author thought about performing institution design using the engineering design method. We all know that in architecture, there are drawings; in electrical works, circuit diagrams; and in mechanical design, mechanical drawings. With the aid of these drawings, project designers may go about their work in both an intuitive and an objective/scientific manner with high operability, with each improvement made to the design in question grounded and practical. In contrast, in the fields of economics and management studies, despite the popularity of institution research, the study results obtained are often recommendations that are deemed to present "inadequate supervision," "insufficient impact," or which are "excessive intervention." Hence, despite the popularity of institution research over the years, the practical applications of such research remain far and few in between.

The key reason for this situation is the lack of common institution design tools and techniques, with the vast majority of studies relying on personal experience and thinking. This means that the resulting studies lack depth and are flawed. Without the use of an overall institution diagram annotated with clear symbols, an over-reliance on intuition and experience can make it difficult to observe the multivariate factors in, and complex structure of, an institution. Without the use of a mathematical model, a precise comparison and consideration of effects and costs of various institutions cannot be performed. A variety of solutions have been proposed for similar problems plaguing the institution, and this is precisely because these solutions have been forged in individual, personal experiences.

Years have come and gone (the article published in *Theories and Practice of Systems Engineering* in 1995 was written in 1993), and in these 20 years, the author has gone through successive attempts and failures in his efforts to develop a viable set of rules and a symbol system for institution design drawings. Finally, in July 2013, he achieved a key breakthrough, successfully developing a set of symbols

and rules for institution design. As the original name was too long and unwieldy, and in order to commemorate the two decades the author had spent on the subject (he started at the age of 40, and is now aged…), the author decided to term this diagram the "Sun Diagram."

Standing in contrast to the "Sun Diagram" is the "Behavior Reward Function," a simple function that conforms to the law of diminishing marginal rewards and which was discovered at the same time. The use of the function in conjunction with the Sun Diagram makes for a simple and incisive analysis of various institutions. When reading this volume, the reader will soon discover that the author seeks to keep mathematical model used clear and simple without redundancies.

Through analysis with the "Sun Diagram" and the "Behavior Reward Function," the author has discovered that there are five fundamental institution structures across all of human society from which derivatives and adaptations arise.

This volume provides practical methods of institution design and analysis, including the "Sun Diagram" that reflects institution structures and institution components' configuration. These methods all have tremendous practical applications.

This volume also includes a number of significant findings made by the author which come with practical applications.

When there is production by an organization, if production behavior does not produce any externalities, independent operation and management should be prioritized ahead of other models as enterprise production will reach spontaneous equilibrium of optimal scale of production under such model, without needing any management costs.

Where there are positive externalities from production behavior, the spontaneous equilibrium of the enterprise's production scale will decrease and fail to reach optimal scale. One key finding of this volume is that the subsidy mechanism can lead to the spontaneous growth of production scale where production scale has positive externalities. Therefore, if we want the enterprise to spontaneously grow its production scale, we would need to adopt a cost subsidy policy for the enterprise. Production with positive externalities includes production activities like the construction of basic infrastructure such as reservoirs, roads, and bridges.

Where there are negative externalities from production behavior, the spontaneous equilibrium of the enterprise's production scale will increase and cause the drying up of resources. With regard to this issue, the analysis result in this volume shows that tax mechanisms can be used to shrink production scale for production with negative externalities, so that resources may be better conserved. Hence, a high-tax mechanism needs to be put in place to shrink enterprise production scale in order to better protect the environment and to conserve resources. Such production activities come with negative externalities such as the consumption of vast amounts of water, mineral resources, fishery and forestry resources, or atmospheric, water, and soil pollution. One discovery made here is that the importance of tax collection lies not in the redistribution of citizens' income but in reducing the overconsumption of various resources. This explains the puzzle of why countries with high tax rates tend to have better-protected natural resources and environments.

Therefore, from the perspective of management by the state, when production behavior presents either positive or negative externalities, either the free-market mechanism of independent operation by each enterprise be abandoned in favor of a centralized management mechanism or the free-market mechanism be retained with supplementation by policies.

The significance of this conclusion is that it has proved, in theory, the key approach to improving the supply of public goods and protecting our earth effectively lies in the institution.

Around the world, in some countries, production behavior that presents externalities is governed by centralized management, while in others, the free market is allowed to operate with regulatory policies in place. However, with the growing scale and increasing complexity of social institutions, centralized management is becoming a less attractive option by the day, for reasons of the need to communicate across multiple levels, significant distortion, and loss of fidelity. In a management institution where the free market is supplemented with regulatory policies, the information chain of an autonomously run enterprise is much shorter, which means that the enterprise is more responsive. If the state implements its regulatory policies effectively, the production scale of the enterprise will grow swiftly with self-regulation and reach an equilibrium point with optimal social benefits. Relatively speaking, this would mean a more successful outcome.

For some enterprises, different externalities associated with production emerge under various circumstances. For instance, for enterprises in the cultural industry, such as those in publishing and journalism, if their product aids the development of society—such as when they publish motivational books or volumes with scientific content—their production comes with positive externalities. If their products bring harm to society, such as when such products promote superstition or contain pornographic content, then their production presents negative externalities. Based on the analyses in this volume, if these enterprises are allowed to run their operations with complete freedom as in the free-market model, then production behavior with positive externalities will have a smaller production scale due to lack of effort, while on the other hand production with negative externalities will tend to become greater as a higher level of effort is applied. Therefore, for enterprises of this type, either centralized management should be applied or governmental subsidies should be applied to production with positive externalities to promote the scale of such production, while high taxes be levied on production that presents negative externalities to reduce profits of such production.

For research and development ("R&D") behavior that inherently presents positive externalities, as R&D outcomes can easily be copied (with an externality produced, i.e., benefits for the copier), the equilibrium point of self-motivated effort is comparatively lower. Hence, currently in various countries, there are measures that protect innovators, i.e., patent rights measures. This type of institution can lead to a lack of visible externalities from R&D behavior within a certain period (with the copying of R&D results in this period being illegal), thereby optimizing the effort equilibrium for R&D behavior. The problem is, if patent rights institutions are

not run effectively and copying behavior cannot be hence curbed effectively, individuals would then lose their enthusiasm for R&D and innovation.

As for the competitive behavior often seen in the society, the finding of this volume is as opposed to the common assumption that the more crowded the field, the fiercer the competition, and the competition is fiercest when there are only two units competing with each other. This means that in order to tackle vicious competition such as the malicious suppression of competitors, the right approach would be to allow more enterprises into the market rather than to reduce the number of players in the market. This conclusion explains why two persons pitted against each other in a competition would engage in rivalry greater than they would in a multiplayer competition. In international relations, this explanation also provides the theoretical ground for why a multipolar world would be more peaceful and stable than a bipolar world. Indeed, we can all remember how two world superpowers had nearly brought mankind to the brink of nuclear war during the Cold War period.

These findings may be harnessed in institution design to tackle key challenges that humankind faces today, such as environmental pollution, overconsumption of natural resources, carbon emissions, peace and stability issues, and stagnating productivity levels.

Further, the engineering method for institution design proposed in this volume allows the process of institution design to become more like the process of engineering design, with a choice of graphics and institution components for the designer as well as the ability to calculate and compare institution effects for the optimization of design plans. This makes for institution designs that offer greater operability and which are more practical, making significant impact from the scientific method on management practice.

Finally, in the area of fundamental theory, this volume proposes the three conditions necessary for any sort of behavior: behavioral utility (positive utility), behavioral resources, and behavioral opportunity. All the three conditions must be present at the same time. Clearly, this is significantly different from traditional theories of economics where only behavioral utility is considered. In addition, the proposal of the concept of exchange utility has also provided an effective method for considering the scale of benefits of different types.

For over 20 years, I have been indebted to the generosity of a number of R&D funders (the contents in this volume include the outcomes of what have been funded, but not in their entirety), who are the National Natural Science Foundation of China (71171134, 70871080, 70471066, 70271005); the Shanghai Key Basic Research Project (03JC14054); the Shanghai Social Science Planning Project (2011BGL006); the Innovation Project of Shanghai Municipal Education Commission (01F06, 11ZS138); the Shanghai Top-Tier Disciplines Project (S1201YLXK, A14006); the Doctoral Fund of Ministry of Education of China (20070252002); and the Management Science and Engineering of Shanghai Municipal Plateau Discipline.

November 2014 Shaorong Sun

Contents

Chapter 1
General Principles of Institution Design

1.1 Institutions as a Key Tool in Behavior Management

1.1.1 Culture and Institutions as Two Fundamental Tools in Behavior Management

Mankind has a long history of thousands and thousands of years. In fact, the history of mankind is the history of human behaviors. Labor, learning, research, building, war, etc.: all these activities are human behavior in progress.

As society accumulates ever-greater amounts of wealth and as living conditions improve, mankind has become more diverse as it grows. Today, we have a global population of 7 billion persons. However, human behaviors have been kept coordinated and orderly despite the proliferating number of people and activities. How was this achieved?

This was achieved through behavior management.

There are two basic tools in behavior management: culture and institutions.

1.1.1.1 Culture as a "Soft Limit" on Behavior

Culture is a key tool in behavior management. Traditional customs, ways of living, behavior standards, values and beliefs, etc., all have an impact on standardizing and limiting human behavior.

In general, culture (take note that "culture" as used in this volume is a narrowly defined term that is related only to behavior and behavioral tendencies) is an understanding of the individual on what kind of behavior is appropriate in a given situation. It exists as a kind of consciousness. Under various circumstances, culture can lead to an individual choosing to act in a certain way and avoid certain behaviors. For instance, for behaviors such as taking care of one's family and the

S. Sun, *Five Basic Institution Structures and Institutional Economics*,
DOI 10.1007/978-981-10-0344-8_1

elderly, culture plays a key role. The impact of culture in guiding human behavior is comprehensive and thorough, and selected behavior takes place on a voluntary basis. However, when it comes to changing human behavior, culture's impact takes place gradually, over a period of time, and ambiguously. This is because culture itself evolves gradually and its affect can be ambiguous. If we are to rely on culture alone to correct undesirable behavior, it may take a very long time before results are seen. Therefore, we can say that culture acts as a "soft limit" on human behavior.

In addition, despite the fundamental impact of culture on an individual's behavioral choices, it is not the be all and end all. In fact, it is often limited in impact and scope. For instance, individuals can be very generous with their own kin and are willing to drop their personal interests to help their family members. Clearly, culture has had a large impact here. However, when their interests are in conflict with individuals outside of their families, they tend to show the selfish side of their personalities. At this juncture, if we are to depend solely on culture to guide our choices in so as to fairly coordinate benefits between members of society we would find ourselves in quite a tough situation. Here, the effect of the institution comes into play.

1.1.1.2 Institutions as "Hard Limits" on Behavior

Another key tool in behavior management is the institution. The institution in the context of the state would be state laws; and in the work unit, rules and regulations. A management institution is formed when behavioral standards (such as what should be done, what should not be done, and how to act in specific circumstances) and rules regulating behavior (i.e., rules regarding the behavioral rewards, resources, opportunities and costs, etc., to be granted to the individual under management with reference to a certain behavior) are concretized into institution rules and the management institution when the implementers of such institution rules are named. When individuals make behavioral choices with a management institution in place, they will clearly feel the limiting effect of the institution. Therefore, we can say that institutions act as a "hard limit" on human behavior.

To illustrate the importance of the institution in behavior management, we now have a story about how a group of seven persons, who were to share a bucket of gruel a day, dealt with the problem of gruel distribution at mealtime.[1] However, we should note that there was not much gruel and a little short of enough for everyone (Wang 2004).

In the beginning, it was decided within the group that each person would take turns to distribute the gruel. The result was: the only time any person would feel sated would be the day he was put in charge of dividing up the gruel.

[1]This story was excerpted from Wang (2004). Changes have been made for the purposes of this volume.

Later on, a person who everyone had held to be "of superior moral values" was nominated to be solely responsible for the task. However, as they say, power corrupts, and soon everyone else tried to bribe him. At the beginning, the nominated individual did well, standing by the belief that he should "live up to everyone's trust" to "distribute the gruel fairly". Over time, however, he was unable to keep to the "moral baseline" and the "unequal distribution of gruel" became a recurrent event.

To restrict the powers of the "gruel distributor", a council meeting mechanism was proposed. So, a three man "Gruel Distribution Council" was established alongside a four-man "Gruel Evaluation Council". The two Councils ended up in an endless feud. The problem is: these two "Councils" were not feuding with each other on the matter of the "equal distribution of gruel". Rather, the in-fighting was due to vested interests on each side. The feud became very intense as both factions competed for benefits and vanity. Numerous drafts of the "Gruel Distribution Plan" were written, and discussions followed discussions. Eventually, everyone became famished from their struggles. By the time the "Gruel Distribution Plan" was approved, the gruel had gone cold, resulting in some of the members suffering stomach problems eventually.

After a long process of seeking the right solution the group finally settled on a new "gruel distribution institution": "the one who distributes the gruel will get his share last". In other words, the one responsible for distributing the gruel will get the last bowl of gruel left after everyone else had made their selections. This way, the person in charge of gruel distribution would work to avoid getting the bowl with the least amount of gruel by making sure that the gruel is equally divided across every bowl. Thereafter, there was no more conflict within the group, and everyone lived happily ever after.

Here, we see that different institutions can generate vastly different outcomes with the same problem.

While the example above fictional, we now look at a real-life case.

Prior to the opening up and reform of the Chinese economy, the communal nature of resource distribution meant that there was no incentive for one to work harder. As a result, there was very little motivation and enthusiasm for work among individuals. Residents in Fengyang County in Anhui Province had annual rations of 150 kg per capita and earned an annual per capita income of just RMB50. As production output remained at a stagnant level, the leadership of the production team was changed frequently in Xiaogang Village in Fengyang County. Eventually, all 17 males with labor capacity had taken on the head or deputy head positions at the team. As the replacement of team leadership typically occurred in the fall after low production had been tallied, the locals had even a saying for the phenomenon: "when the numbers are low, the heads will roll".

In spring 1978, the Xinjie Commune in Tianchang County, Anhui Province met with a devastating drought that had threatened to kill off the cotton saplings in the fields. It was an unprecedented disaster in the making for all concerned. To ensure its survival, the commune had no choice but to abolish communal production, making it known that each household would be awarded or penalized based on its

level of production. With the rule "the more you work, the more you get" explicitly in place, members of the commune worked hard to fight the drought, even reserving the water used from washing their faces for the watering of the cotton plants. In the end, despite the drought, Xinjie's cotton production levels increased by a hefty 90 % over the previous year!

It is clear that the "design of an effective institution" is critical to behavior management. Indeed, institution design is fundamental and presents a broad range of applications. Wherever there is an organized group of persons, there is an institution.

1.1.2 Culture and Institutions Must Be Taken Care of at the Same Time

Culture and institutions complement each other in modern behavior management, and neither one of these two should be neglected.

The two share similarities as well as differ from each other in terms of their respective characteristics.

The similarity between the two are, from the perspective of their respective structural elements, both are comprised of two types of elements: the first is a set of specific behavioral rules which stipulate what ought to be done under specific circumstances, i.e., the appropriate actions one should take in various situations. Second, within both culture and institutions there are "implementers" who enforce the respective rules or codes.

The difference between the two is, from the perspective of behavioral rules, the rules presented by culture tend to be more ambiguous and presented in terms of general principles, while the institution behavioral rules tend to be more concrete and specific. Another difference is, from the perspective of the "implementer" whose job is to limit behavior, in culture the role is uncertain: sometimes it manifests as "self-awareness" on the part of the actor while other times the actor is "forced to comply" due to pressure from others. Moreover, "others" can also refer to a variety of players: one's friends, family members, neighbors, etc. In contrast, the "implementer" in an institution is specifically designated as such, and often there is a dedicated entity designed to induce individuals to comply with the behavioral rules stipulated within the institution.

From the perspective of behavior management, the two elements complement each other.

The impact of culture in behavior management is then it tends to have a comprehensive and total effect and does not require the use of supervisory and implementation measures. However, when it comes to changing human behavior, culture's impact takes place gradually and ambiguously. Moreover, culture can present high "inertia", which means that once the attitude or belief in question has

been formed it cannot be changed easily. This means that it is difficult to apply culture immediately in behavior management for specific objectives.

Using institutions in behavior management can yield quick results with clear requirements in place. However, institutions that lack the support of prevailing culture tend to be less effective.

First, culture affects one's opinion of a given behavior. For instance, with a "institution designed to curb corruption", the institution designer would often make use of "monitoring by the public" and "tip-offs from the public" to observe corrupt behavior by officials which may then be tackled head-on. However, if everyone were to believe that "it is natural for an official to try and make more money", then their attitudes towards corruption would be less strident. At this juncture, depending on "tip-offs from the public" to help observe corrupt behavior would also become a lot less effective.

The same institution can produce different effects in different cultural environments. For example, an election institution would work very well in some places where residents are able to "vote clean" to elect a candidate who is "both morally upright and talented" to work for the people. However, in other areas, "vote-buying" behavior is commonplace. If the cultural environment were positive, however, then everyone would "reject vote-buying". Here, the outcome of an election institution would be positive. In contrast, if everyone were to believe that "the man who can give me the most benefits is the man for me", then "vote-buying" would become rampant. In such a situation, so-called "democratic elections" would be no better off than the "pre-screening of candidates by higher authorities".

Second, we can find flaws with almost all institutions. It is extremely challenging to design a flawless institution. This is why the institution can only work effectively in conjunction with cultural cues. As for "behavior that is hidden from the view of behavior managers" or "behavior that can only be left untouched", the institution is powerless. In reality, it is precisely because a positive culture can have a management effect on "behavior that is hidden from the view of behavior managers" or "behavior that can only be left untouched" that individuals are able to comply with institution requirements in a "self-aware" manner even in "situations where the institution cannot manage the situation at hand".

Third, the choice of institution to be implemented often depends on cultural tendencies. This means that without cultural support an institution will not succeed in its aims.

On the other hand, a mere dependence on culture to manage behavior will also not yield good results. Even if an extremely positive and deeply embedded culture is in place and individuals are able to comply with cultural norms in a self-aware manner, the ambiguity of culture's impact on behavior management means that it is still difficult to ensure that everyone's actions are coordinated and hence stay orderly. For example, without the use of traffic lights in traffic management, problem would arise at an intersection even if drivers all give way to each other: traffic would not be moving because everyone is just waiting on each other. This type of traffic 'movement' remains inefficient.

Therefore, we can say that culture "nourishes" behavior, while the institution determines behavior. The two must complement each other to ensure the orderly behavior of society.

1.2 Basic Institution Elements and the Task of Institution Design

1.2.1 The Two Key Elements of an Institution: Institution Rules and Institution Implementers

Institution rules and institution implementers are the two critical elements in an institution. The institution is not complete if either element is missing.

Perhaps some would ask: what exactly is an "institution"? Where exactly is the "institution"? Where can we "see the institution" in action?

To determine the existence of a given institution, we must first check to see if corresponding institution rules exist. Then, we must determine if the institution implementer exists.

Like culture, institution rules exist in the minds of individuals under management. Or rather, they exist in the consciousness of these individuals as management tools created in the thoughts of man.

For instance, in social organizations there is the superior-subordinate relationship, with the subordinate reporting to the superior on unusual matters that have arisen as well as taking instructions from the superior. These conventions have emerged because individuals recognize that they need to "seek advice from a superior when unusual circumstances arise". This consciousness is actually the recognition of institution rules within the organization.

It is often said that the "institution" lies in organizational rules, notifications or provisions. Indeed, in order to standardize the recognition of institution rules between a number of individuals, very often institution rules must be clearly and explicitly stated. This is how various regulations and mechanisms such as "notifications" and "guidelines" have come about. However, these regulations and mechanisms are but an external form of the expression of institution rules. The purpose of all these "explicitly stated items" is to let individuals "know about", and "remember", institution rules. Fundamentally speaking, institution rules can only work if they exist in the minds of those being managed.

The stories of "gruel distribution" and the drought-hit cotton producing commune that we have looked at both reflect the importance of the institution.

The institution implementer consists of various management entities and corresponding equipment, with the former the human resource component of the implementer and the latter, the material component. Management departments within a corporation and various agencies within government act as the human resource portion of institution implementers. Equipment required in the course of institution implementation constitutes the material part of institution implementers.

Later on in this volume, we will see that with the use of the Sun Diagram to conduct institution analysis, there are only two main elements in an institution: institution structure and institution components. The structure and components together determine institution parameters. By calculating the parameters of the institution, we can get the indicators of institution effect.

This chapter divides the institution into institution rules and institution implementers. If we are to consider the relationship between the two, we can say that institution rules determine institution structure and at times, the performance of institution components (at this juncture, certain institution rules actually act as the internal rules of institution components with many institution components in themselves "mini institutions") while institution implementers (both agencies and equipment) determine the properties of institution components.

In reality, only when institution rules are reasonable and when implementers have the requisite implementation capability, will the institution be effective.

Example 1.1 The history of the development of traffic management equipment

The institution consists of institution rules and institution implementers. Currently, in various countries around the world, traffic management institutions have been implemented well mainly due to the effectiveness of institution implementers. Apart from traffic policemen, the development of the relevant equipment has also made this possible.

First of all, there has been significant improvement in the equipment that conveys traffic commands. Very early on, the traffic policeman was required to direct traffic using hand signals. However, the visibility of such signals was limited and differences between different signals were limited. As a result, sometimes drivers were unable to read the signals given correctly. In 1868, the world's first traffic light was put into use. A policeman was placed at a junction to switch between two gas lights: a red-colored one for "Stop", and green for "Go". For a long time after that, electric traffic lights were switched manually by policemen standing at junctions. Today, the switching between the red and green lamps is effected through a computer program. When special circumstances arise, duty policemen can also remotely control the switching of signals on-site.

Second, observers meant to observe any breach of traffic regulations have also been improved tremendously. In the past, traffic policemen could only perform visual observation of the breach of traffic regulations by various drivers. This approach tends to present human errors and testimony can be difficult to defend in a court of law. Moreover, many breaches can be hard to determine (for instance speeding, which cannot be determined by the naked eye alone). Today, with the aid of cameras, we can now have records of vehicle speed, license plate numbers, travel routes, whether the vehicle has run a red light, etc. We see that improvement in traffic signaling equipment has made the traffic light institution a highly effective one (Fig. 1.1).

Fig. 1.1 Observer of traffic behavior in breach of institution rules (equipment part): the camera

Example 1.2 Institution rules and institution implementers in China's "gaokao" exam institution

China's *gaokao* or standardized university admissions examination institution has decided the futures of generations and generations of young adults. The institution is similarly comprised of institution rules and institution implementers.

The *gaokao* institution is an examination and admission institution with the institution implementing agency being the exam management authorities and invigilating teams in various provinces while the exam registration, exam results query and university admissions online institutions constitute implementer equipment. There are also camera institutions acting as observers of exams in progress and jamming equipment designed to cut off electronic signals that could be used for cheating, etc. These rules, teams and equipment ensure the fairness of the university admissions institution.

Example 1.3 The wage institution

In a particular workplace, the wage institution is as follows: total wage = number of days employee is at work × job designation coefficient. The implementing agencies and personnel here are the Human Resource Department of the company and the office supervisor respectively. Of the two, the latter is responsible for recording the attendance of every employee daily. (In Chap. 7 we will see that the role of the office supervisor is actually that of the institution "observer") while the Human Resource Department is responsible for the preparation of wage tally documents.

Soon after, problems began to appear: some workers complained that the office supervisor was biased towards certain employees, who were observed to be in the office "on time" despite being terribly late. On the other hand, workers who the supervisor did not get along with were marked down as being late for half an hour even if they had been late for just a couple of minutes. As attendance statistics were released only at the end of the month, workers who had felt "wronged" were unable to raise effective complaints due to the length of time that had passed. The atmosphere in the workplace worsened as a result.

To resolve this issue, it was decided that an "automatic clock-in institution" would be implemented. Within such an institution, workers would clock in using magnetized cards issued to each individual with their clock-in times automatically recorded. The "human bias of attendance records" was thus eliminated.

However, sometime later, a new problem emerged: some tardy workers would hand their cards over to colleagues who would typically arrive earlier for clocking in on their behalf. In other words, there was "fraudulent clocking in" with the result of "falsified attendance records".

Recently, with technological advancement certain clocking institutions are now capable of reading human fingerprints and other biometric data. The phenomenon of "fraudulent clocking in" has become history as a result.

This example illustrates the importance of equipment in an institution, and shows how the tendency in past studies to neglect the impact of improved equipment on institution effect should be corrected.

1.2.2 The Task of the Designer: Optimizing Institution Structure or Deploying Institution Components of Even Stronger Performance

The institution is a structure that comprises institution components linked together in a certain relationship. Typically, different institution structures or institutions of a similar structure but comprising institution components of different characteristics would also differ in terms of institution effect. Therefore, institution designers have two key tasks: to optimize the institution structure or to deploy institution components of even stronger performance.

During institution design, first we must perform a diagnostic examination of the existing institution to determine the reason for the poor performance of the institution. Is it because the institution structure was poorly designed, or is it because of the poor performance of institution components?

If the poor performance of the institution is due to poor structural design, then a re-design of the institution structure would be in order with the aid of the Sun Diagram in institution analysis. The Sun Diagram is a diagram that is similar to the "circuit diagram" in the electronics field and which consists of a symbolic structure.

It serves to provide a description of institution structure. Specific details on the Sun Diagram, such as the symbols used and rules governing the diagram, will be discussed in Chap. 6.

Like the electric components in a circuit diagram, the institution component must exist in reality. In other words, we can only choose institution components that perform well, and which exist in reality, in our institutions. Of course, institution components themselves can also be designed, but this sort of design must be based on viability in practice (e.g., ensuring that certain resources necessary for the functioning of the component in question are in place). Therefore, a collation and review of existing institution components and a comparison of their respective performances together with the design of better-performing institution components also constitute an important task in institution design.

It should be noted that in many situations institution components are actually "mini institutions" that in turn make up a "big institution". Hence, institution components are often also comprised of rules and implementing agencies/equipment. Very often, agencies and equipment alone do not constitute an "institution component" in the true sense of the term. For instance, "wage levels" act as a behavior promoter within the institution. For a behavior management institution, it is an institution component. However, at the same time the "wage levels" mechanism is in itself an institution comprised of rules, implementers and equipment.

We shall look at commonly used institution components in Chaps. 3, 4 and 5 of this volume.

1.3 The Hypothesis of the Rational Man and the Repeatability of Behavior Within an Institution

How individuals under management act within an institution, and whether they act rationally or irrationally, are questions that institution design must address.

1.3.1 Rational and Irrational Man

The hypothesis of the rational man is a fundamental premise of institution design.

This hypothesis refers to the idea that man is able to correctly judge the actual utility of various solutions or options and is also able to, based on the prior judgment, choose the behavior with the greatest utility.

The vast majority of theories in economics and management are premised on the hypothesis of the rational man. Take for example the theory of market equilibrium and Hardin's idea of the "tragedy of the commons". However, ever since the theory of "behavioral finance" by Kahneman et al. won the Nobel Prize for Economics, the hypothesis of the irrational actor has also gained currency.

The so-called "hypothesis of irrational man", simplistically speaking, is the idea that man is unable to correctly determine the actual utility of various behaviors, which is why they "often make the wrong choices". A series of experiments targeting the economic behavior of college students support the theory of irrational man by Kahneman et al. Through these experiments, it was found that oftentimes economic choices made by individuals are determined by the way such choices have been presented rather than the actual utility of various choices. Behavioral choices made by individuals differed widely and sometimes the same individual would make different choices for the same issue over time.

1.3.2 The Repeatability of Behavior Within an Institution Makes the Hypothesis of the Rational Man the Basis for Institution Design

However, in the opinion of this author, although irrational behavior can be found everywhere in everyday life, it actually occurs with certain conditions. In other words, it is not human-driven. In reality, irrational behavior arises in most cases due to insufficient decision-making time, the lack of experience or access to imperfect information, all of which then lead to the inability of the individual to make the correct judgments. In contrast, when given enough time for decision-making, or when the individual has had experience in dealing with a similar situation, or when the individual has access to perfect information, in most cases individuals are able to express strong rationality by selecting the behavior with the greatest utility.

And the key difference between the institution and specific "orders from above" or "administrative intervention" is, the institution is broadly applicable and stable. Institutions are designed for social roles rather than for a specific individual. Institutions are designed to shape behavior that is continuously being repeated over time rather than a singular instance of action. For the individuals targeted for management, as long as they repeat their behaviors over time, they will eventually move from the lack of experience to being experienced, and move from a grasp of imperfect information to the possession of perfect information. Therefore, we see that individuals within an institution would have a fundamentally accurate reading of behavioral utility. Individuals under institution management are by nature rational rather than irrational. Thus, the behavioral choices made by those under management within a given institution are determined and can be predicted. Changes to institution structure and parameters can lead to a change in the behavior of individuals targeted for management. Herein lies the significance of institution design.

It must be noted that we can only say that in general the individuals managed within an institution are rational; however, we cannot claim that all of them are rational. In an actual group of individuals targeted for behavior management within an institution, the rational and irrational understanding of the issue at hand would

vary certainly from person to person. It is precisely this reason that the analysis of the impact of the institution on individual behavior using the Sun Diagram and mathematical models in Chaps. 7, 9 and 10 are premised purely on the hypothesis of the rational man. However, in Chap. 8, when we analyze the impact of the institution on the group, we will use the concept of "behavior probability" as a way of expressing the rationality and irrationality within a group. The conclusions derived will be all of a probabilistic nature rather than a conclusive nature.

1.3.3 The Difference Between a Comprehensive Concept of the Rational Man and the Classic Concept of the Rational Man

However, this assumption of human nature in this volume is that of the "holistic rational man" which differs from the classic concept of the "rational man" in economics.

The "rational man" in classic economics, often also known as the "economic man", refers to an individual whose criterion for judgment is the degree of economic benefit. In other words, the classic "rational man" is only interested in the so-called "maximization of economic benefits".

The "rational man" as defined in this volume is an individual whose judgment criteria are of a more varied and holistic nature. Here, the individual is interested in "economic benefits" as well as "social benefits" such as a higher social standing and a better reputation and "psychological comfort". All these must be considered when determining "behavioral utility". In fact, often in behavior management, certain non-economic incentives and penalties are also used. Such mechanisms include workplace promotions, the demotion of personnel who have made severe mistakes, the commendation of personnel who have acted well, and the reporting and criticism of personnel who have behaved inappropriately, etc.

Therefore, more accurately speaking, the rational man as assumed in this volume is a "holistic rational man" who is different from the rational man of classic economics in that he is interested in both economic and non-economic benefits.

1.4 The Difference Between Gaming Within an Institution and Gaming in General

When the individual selects the behavior with the greatest effect to himself within the institution, he is actually engaging in trying to "game" the institution. However, in contrast to the general idea of "gaming", the gaming involving the individual within an institution has several unique features.

1.4.1 Gaming Within an Institution Is Strategic While Individual "Gaming" Behavior Is Simple

Within the institution, when the individual is trying to "game" the institution, institution behavior is strategic with a kind of "response rule" expressed. This kind of response rule is structured as: "if the individual engages in behavior a, then the institution shall respond with behavior A", and "if the individual engages in behavior b, then the institution shall respond with behavior B". For instance, the strategic rule of the wage institution at a certain company is: "if the candidate is armed with strong academic qualifications, he shall be given a high wage; if he is not armed with strong academic qualifications, he shall be given a lower wage." This way, the individual may, based on the institution strategy, select his preferred strategy.

For individuals targeted for management within an institution, the "gaming" that they undertake is simple. To take the above example, with such a strategic rule in place, the best choices for the individual would be to "study more" in order to enhance his qualifications.

1.4.2 Commitment in "Announcement" and Commitment to "Realization": How the Existence of the Institution Is Expressed

From the perspective of those targeted for management, gaming the institution is expressed as two kinds of commitments: the commitment made in the "announcement" [of institution rules] and the commitment to [rule] "realization", as in the statement "if you choose behavior X, I (i.e., the institution) will choose behavior Y". For example, a country's traffic regulations may stipulate that "running the red light" is not allowed, with the public informed that "drivers caught running the red light will receive six demerit points on their driver's license and will be fined 600 yuan". At the same time, it would also be made clear that "if the driver does not run a red light, he would be allowed to proceed as normal". This way, the individual will—based on the institution commitments expressed—generally opt for "not running the red light".

Of course, to make sure that individuals comply with institution management, institution promises must be "believable" or convincing. Generally speaking, the believability of institution commitments can come from two sources: the first source would be the adequacy of institution conditions, such as the implementation capability of institution teams and equipment. The second source would be an experiential knowledge of the institution, i.e., if, based on past experiences, the institution will fulfill its stated commitments.

1.4.3 Individual Makes Behavioral Choice Before Institution Fulfills Commitment

Within the institution, individuals targeted for behavioral management would first observe institution commitments before opting for one behavior or another following a consideration of all relevant factors. On the other hand, the institution observes the behavior of the individual and then selects its corresponding behavior based on pre-stated rules (i.e., commitments) in order to fulfill the relevant commitments.

Generally speaking, in this kind of gaming between the individual and the institution, the response rules of the individual's "opponent" (i.e., the institution) tend to be more stable, and this stability of institution response rules is the very expression of the existence of the institution. Under such an institution environment, the individual may maximize utility for himself through behavior choice. Clearly, this sort of game where only one party is truly free to make behavioral choices is different from the kind of game that goes on in general where both parties are free to select their respective behaviors in a "behavior set".

As these are the characteristics of the game within an institution, we can see that the only time we have flexibility with the selection of the institution game is during the process of institution design. Once the institution has been put in place, its response rules may not be changed at a whim. Therefore, much care must be taken in institution design. Before an institution design proposal can be confirmed, the scientific prediction of individuals' behavior must be performed. Only then can an effective institution be put in place to ensure the consistency of selected behavior with institution aims.

Reference

Wang X (2004) Lessons from the story of gruel distribution. J Hangzhou Financ Training Inst 1:54

Chapter 2
The Three Conditions for Behavior, and the Five Behavior Management Measures

2.1 The Three Necessary Conditions for Behavior

The three necessary conditions for the occurrence of a given behavior are behavioral utility (positive utility, since we are assuming that the actor is rational), behavioral resource, and behavioral opportunity. All three conditions are necessary for the occurrence of the behavior in question. Unlike in classic economics which is merely concerned with behavioral utility, a key principle in this volume is that for a given behavior to happen all three conditions must exist.

We can express the relationship between the three necessary functions as: $e = f(u \wedge res \wedge opp)$. Here, e stands for the behavior effort level, $f(\cdot)$ the incremental function, u the positive utility of said behavior, *res* the behavioral resources, and *opp* behavioral opportunity. $\wedge$ is the logical operator, indicating that the elements to its left and right must exist or that one should take the smallest value of the elements to its left and right.

2.1.1 Behavioral Utility

Roughly speaking, the term "behavioral utility" refers to the "benefits" arising following the action and which meet the needs of the individual. Of course, if certain behaviors bring "drawbacks" to the individual, the behavioral utility here would be "negative".

Without doubt, when an individual has to choose a certain action among several options, behavioral utility becomes a critical factor in consideration. Man hunts for food when he is hungry, and seeks additional layers of clothing when he is feeling cold. Both these instances are examples of how behavioral utility would affect one's priorities.

S. Sun, *Five Basic Institution Structures and Institutional Economics*,
DOI 10.1007/978-981-10-0344-8_2

In this volume, we assume that the actor is risk-neutral, and this behavioral utility can be directly expressed as benefits (such as economic benefits). Under such circumstances, behavioral utility u can be regarded as the difference between two factors: behavioral reward r and behavioral cost c, i.e., $u = r - c$.

Most, if not all, studies in the areas of management science and economies on the subject are solely concerned with utility as a cause of behavior. For instance, Abraham Maslow has proposed the theory of the Hierarchy of Needs, which states that only unsatisfied needs can affect the behaviors of the individual. For the theory, he proposed five "steps" or levels of needs: biological needs, security needs, social or emotional needs, the need for respect, and the need for self-actualization. Frederick Herzberg's Motivation-Hygiene Theory states that there are two key factors that affect human behavior: hygiene and motivation. In this theory, the hygiene factor can eliminate "dissatisfaction" in individuals while the motivation factor can bring "satisfaction" to individuals. The Expectancy Theory of Victor Vroom states that only when the individual anticipates an attractive outcome to a certain behavior would he opt for such behavior. In his Reinforcement Theory, B.F. Skinner states that human behavior is a function of stimulation. If the stimulation brought on by a certain action is favorable, then this action will surface repeatedly. On the other hand, if the stimulation is unfavorable, the tendency for this behavior to occur again would be diminished.

As for the latest developments in the field of behavior management, Leonid Hurwicz et al., have proposed theories of mechanism design including the principle of incentive compatibility, which is also based on behavioral utility.

In management practice, the idea of behavioral utility has also received much attention. For instance, in studies of the subject of remuneration as well as in remuneration plan designs, behavioral utility is regarded as the way to promote hard work by the employee (i.e., to encourage the worker to select the behavior "working hard").

From the above, we can see that in traditional studies of behavior management, behavioral utility has been only regarded as the variable capable of influencing behavioral outcome.

However, to look at the factors that affect human behavior holistically, apart from the traditional "behavioral utility" (which has been studied the most) we must also look at the factors of behavioral resources and behavioral opportunity.

2.1.2 Behavioral Resources

The term "behavioral resources" refers to consumables used up during the behavioral process. For instance, all actions require time to complete. Therefore, time is a resource required for the occurrence of behavior. This is also a key reason why humans cannot engage in too many activities at the same time. To take another example, in research behavior, apart from time, funds are also needed for the

purchase of equipment and materials. Therefore, funds are a commonly used form of behavioral resources. In a modern society of commodities, funds have a certain degree of versatility as many behavioral resources can be exchanged by using funds.

2.1.3 Behavioral Opportunity

Apart from behavioral utility and behavioral resources, behavioral opportunity is another factor that has an impact on the probability that a given behavior would occur. If the actor has objectively opted for a certain behavior or action, and the resources required for the action are adequate, then objectively speaking the probability that the given behavior would occur would be the chance of behavior.

For instance, for a policeman wishing to prove himself by catching burglars, if in his work environment he only comes across individuals who display desirable behavior, then his chances of meeting, and catching burglars would be diminished.

To take another example, during the Cultural Revolution period in China, as the *gaokao* examination institution had been canceled, youth of the right age group at that time had zero possibility of "gaining admission to university".

2.2 The Five Measures of Behavior Management

We can hence see that three conditions are necessary for the actor to undertake any behavior or action. They are: behavioral resources, positive behavioral utility, and behavioral opportunity. Here, behavioral utility comprises behavioral rewards and behavioral costs.

Hence, institution designers can choose from five types of measures when trying to create institutions that impact on individual behaviors: measures that control behavioral resources, measures that control the scale of behavioral rewards, measures that control the extent of behavioral costs, measures that control the availability of behavioral opportunity, and measures that change observation intensity or sensitivity. Of these five types of measures, the first four types are targeted at the three conditions that make behaviors possible (of which the utility condition can be further divided into reward and cost conditions), while the last measure is a generic measure that can used in tandem with any of the other four measures.

We need to note that although these measures are management measures in nature, they are often affected by technological and material conditions. For example, reward measures require managers to have access to a certain amount of monetary and material resources, while behavioral cost measures tend to be associated with technological means.

2.2.1 Reward Measures

2.2.1.1 The Concept of the Reward Measures

Behavioral reward measures are a type of management measure that works by making changes to behavioral rewards so as to influence the actor's proclivity towards a certain behavior. Therefore, with desirable behavior the manager would provide positive rewards while for undesirable behavior, the manager would try to provide negative rewards for the actor.

Behavioral rewards can further be classified as naturally-occurring rewards and rewards granted by the manager. Examples of the former include operating profits and social standing, while examples of the latter include wages, bonuses and fines, job promotions and demotions, etc. The rewards granted by the manager are also known as "managerial rewards".

Since ancient times, the reward measure has been the most commonly used measure in behavior management.

Example 2.1 "High pay to ensure incorruptibility of public servants"

In some countries, to tackle the problem of corruption among public servants, the institution of "high pay" has been adopted with significant pay increases for such workers. This method is an effective reward measure that increases the negative reward associated with corruption. If the official's corrupt behavior were to be detected, not only would he receive administrative punishment he would also lose his high pay. This way, the negative reward from corrupt behavior would be exacerbated, making the option less attractive for officials.

2.2.1.2 Reward Types

Generally speaking, the number of reward measure types equals the number of reward types available. Key reward types are as follows:

(i) Economic-type rewards such as cash rewards, fines, and wage. This type of reward can satisfy individuals' need to improve the quality of their lives. Reward measures targeting this type of rewards are used most commonly.
(ii) Reputational rewards such as commendations, criticism, and job promotions and demotions. This type of reward can satisfy individuals' need for social standing and good reputation.
(iii) Attitudinal and emotional rewards that satisfy individual's needs in terms of their beliefs and attitudes, as well as their emotional well-being.

Attitudinal rewards arise with the pursuit of justice or certain moral values by individuals. Such behaviors include helping the disadvantaged and the punishing of villains. The characteristic of this class of rewards is that "rewards" are felt after the occurrence of certain situations by others (and sometimes also the actor himself).

For example, an individual might feel gratified upon hearing that a villain has been punished. This can be seen as a positive reward brought by the legal institution to the public. When witnessing a disadvantaged person struggling to make ends meet with no help in sight, individuals often feel disturbed. This can be seen as a negative reward caused to the public by the relevant institutions that have not been well-designed or implemented.

To take another example, in a hostage situation, the government or the public may sacrifice their respective interests to stage a rescue, sometimes by offering the hostage-taker certain material benefits in exchange for a release. Such behavior is also the result of attitudinal rewards.

Emotional rewards arise from the strong sense of responsibility and attachment that individuals feel vis-à-vis persons they have a close relationship with. For instance, it is precisely out of a sense of responsibility and attachment that individuals try to obtain as much benefits as possible for their children, subordinates, friends, nation, etc. When individuals see that persons they are closely related to have obtained a certain benefit, they would also feel a positive reward. Conversely, when individuals see that persons they are closely related to have lost certain benefits, they would feel a negative reward. This type of reward also has the feature of making the individual feel the "reward" when there is a change to the circumstances of others.

Similarly, although one often incurs a certain loss due to retaliatory actions taken against one's enemies or competitors, as long as the opposite has taken a beating, the individual would also feel gratified. This is the result of the emotional rewards.

2.2.2 Resource Measures

Resource measures are a type of management measure that works by making changes to resources needed by the actor for committing a certain behavior, so as to have an impact on the probability of behavior occurrence. Therefore, with desirable behavior, the manager would provide more of such resources while for undesirable behavior, the manager would as much as possible try to withhold such resources.

Example 2.2 Provision of research resources by the National Natural Science Foundation of China in order to promote desirable behavior

To promote the desirable behavior that is basic science research, each year the National Natural Science Foundation would provide support to selected research projects. This is effectively a way of providing certain resources to promote basic science research as part of its efforts to enhance the quality of basic science research in China.

Example 2.3 Tackling campus drinking by loading bursary grants directly onto meal cards

A certain college has discovered that some of its students have been drinking on campus, with undesirable consequences such as the frittering away of money on

alcohol, the occurrence of incidents, and negative impact on students' academic performance. Hence, the college has expressly forbidden the drinking of alcohol by students in its rules. However, soon enough the school realized that this rule was not very well enforced. Certain students would buy alcohol as soon as their bursary grants were disbursed, and borrowed from their teachers and fellow students when money ran out. Later on, the school decided to load bursary grants directly onto student meal cards, with these cards only usable in campus cafeterias. Students were also not permitted to cash out any value remaining on these meal cards. As a result, the phenomenon of campus drinking diminished significantly.

In this example, the measure to tackle the problem of campus drinking is effectively a resource measure that has taken away a key resource for such behavior to happen, i.e., cash. As monetary value stored within meal cards could not be used in shops outside of campus cafeterias (which did not sell alcohol), students could no longer purchase alcohol using their bursary grants. This led to students "having the money to eat, but not to buy alcohol", and the problem of "campus drinking" was effectively dealt with.

When using resource-type measures, institution designers must pay attention to the precision of the measure. That is to say, the designer must make sure that resources provided would be used on desirable behavior and not for other purposes.

2.2.3 Behavioral Cost Measures

Behavioral cost measures are a type of management measure that works by making changes to behavioral costs so as to have an impact on the probability of behavior occurrence. Therefore, with a desirable behavior, the manager would try to lower the costs of the behavior while with an undesirable behavior, the manager would try to raise the costs of such behavior. In the case of customs taxes levied on imported goods, the aim is to increase the costs of purchasing imported goods in order to protect the domestic industry.

2.2.4 Opportunity Measures

Opportunity measures are a type of management measure that works by making changes to behavioral opportunity, so as to have an impact on the probability of behavior occurrence. Therefore, with a desirable behavior, the manager would try to provide more opportunities while with an undesirable behavior, the manager would try to eliminate opportunities for such behavior.

The commonly used term "opportunity cost" actually covers two types of situations: first, it refers to the loss caused by the loss of opportunity, such as in the case where one has possibly lost out on a job promotion because he or she had chosen to go back to school. Second, the term can also refer to the impossibility of

re-opting for the behavior due to resource limitations. For example, a stock investor may not have been able to purchase gold at one point due to insufficient funds. This would later cause the investor to miss out on subsequent gains in gold price hike.

Of the aforementioned two situations, the result of the first situation is similar to the result of an opportunity measure, while in the second situation behavior was restricted by a lack of resources, which in essence is similar to the effect of a resource measure.

Opportunity measures are also often used in management practice. For instance, the mandating of community activities for students can cut down the time students otherwise spend at Internet cafes since the opportunity to "hang out at the Internet cafe" has been diminished.

Example 2.4 Tackling the problem of fare cheats

In a certain city, it was found that many bus passengers tended to underpay for their bus fares, thereby causing the loss of operating income for the bus institution. Passengers were able to do that as during ticket purchase at boarding, they would be given different options to choose from at the automated ticket machine in terms of "fare tier". This provided the opportunity for passengers to select fare options that did not commensurate with the actual length of their journeys. Later on, the bus institution adopted a new fare structure of "flat fee of 2 Yuan" no matter the distance the passenger would travel. This eliminated the opportunity for passengers to opt for cheaper fares and hence completely eradicated the undesirable behavior of "cheating on bus fare".

2.2.5 Observational Measures

Observational measures work by enhancing observation intensity or sensitivity vis-à-vis the target behavior in order to increase the probability that said behavior would be affected by the management institution. Such measures are used to encourage actors to opt for desirable behavior and to abandon undesirable behavior.

The observational measure is different from other measures in the sense that it is typically used in tandem with other measures rather than being used singly. In addition, a second difference is: the observational measure has an even closer relationship to equipment.

Example 2.5 Computer technology in aid of observational measures: the example of coal tolls

For many years, Shanxi's highway coal management agencies had been plagued by the problem of bribes made to personnel to facilitate the passage of coal trucks carrying coal for which the requisite tolls have not been paid. This was a problem that caused massive monetary loss. In 1998, this problem was finally resolved with the implementation of a computerized management institution.

This institution comprises of two sub-institutions: an electronic monitoring institution and an electronic currency institution. All vehicles seeking passage are recorded electronically with the weight of the vehicle also recorded and displayed. Coal stations along the highway no longer dealt in cash collection, with payment made using stored-value magnetic cards instead. This effectively eliminated any exchange of cash between station personnel and truck drivers.

The implementation of this institution led to the elimination of various forms of corruption. At the Yangguang Coal Station in Shanxi Province alone, income increased by RMB9.11 million in a single month (Duan 1998).

Commentary:

This is the effect of an enhancement of the observational measure using more advanced equipment, and as a result the undesirable behavior of "coal station personnel being bribed by truck drivers" has been effectively tackled. The problem of corrupt staff allowing the passage of trucks carrying coal without paying tolls was rampant as high-performance observational measures had not been found then. The use of the computerized monitoring institution and magnetic card payment institution has made monitoring of such an undesirable behavior highly sensitive, which means that the relevant personnel can no longer benefit from the abuse of their positions.

This example shows that when seeking effective observational measures, we must not neglect the importance of equipment. Therefore, we should pay attention to the development and use of technological aids.

Example 2.6 The inscription of maker name on bricks: a Ming Dynasty solution against "tofu-dreg" construction (Gao 2011).

In 1372, the fifth year of Hongwu Period of the Ming Dynasty, the head of the Changsha Defense Command Qiu Guang ordered that bricks used to repair the city wall be inscribed with the name of the brick maker and the year of manufacture. This way, if quality issues were later found at the kiln production, it would be easy to track down the person responsible.

Commentary:

This method of inscribing the name of the maker and the year of manufacture onto the brick is a way of realizing an observation measure through the implementation of certain rules on behaviors. This practice has survived into the present day.

2.2.6 The Relationship Diagram of the Three Conditions for Behavior and the Five Behavior Management Measures

We can show the overall relationship between the three conditions necessary for behavior and the five types of management measures using Fig. 2.1.

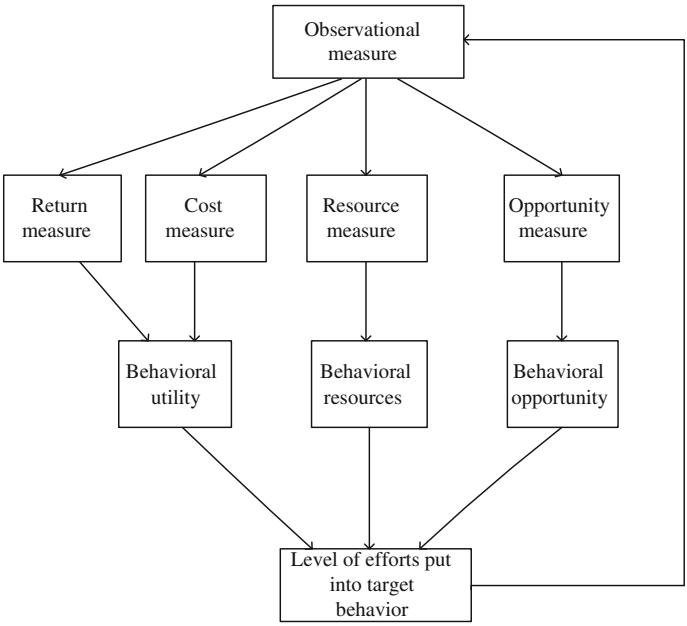

Fig. 2.1 The relationship diagram of the three conditions for behavior and the five behavior management measures

From Fig. 2.1 we can see that the observational measure provides the ground for the operation of other measures. Reward and cost measures can change behavioral utility, while resource measure can have an impact on the scale or availability of behavioral resources. Finally, opportunity measure has an impact on the availability of behavioral opportunity to the individual. The three conditions of behavioral utility, behavioral resources, and behavioral opportunity together determine the chances that a given behavior would occur, as well as the level of efforts put into said behavior by the individual.

2.3 Issues to Consider and Resolve When Using Reward Measures

Compared to the other four measures, the reward measure has been used most commonly in management practice and it also presents in the most number of ways. However, there are also many problems associated with its use. Therefore, in this section we look at the issues to consider and resolve when using reward measures.

2.3.1 The Numerical and Precedence Expectancy of the Reward

As behavioral rewards occur following the completion of the action (i.e., behavior), it means that the actor is effectively making behavioral choices based on his expectations of the rewards a given behavior would bring. Reward measures also work by changing the actor's expectations of rewards from a given behavior, thereby influencing his choice of behaviors.

Therefore, before deploying a certain reward measure, we would need to first to determine the "degree of impact" on the actor by the reward imposed in a precise manner. As such, we would need to understand how actors determine the scale of the reward.

Actors mostly judge the expected reward of a given behavior in one of two ways: numerical expectancy and precedence expectancy.

2.3.1.1 The Numerical Expectancy of the Reward

The numerical expectancy of the reward is a value derived by first multiplying each of the various possible outcomes of a given behavior (often, a single action can result in an array of possible outcomes) with the probability of such an outcome occurring and then adding these sums together. When specific data is available (including the set of possible behavior outcomes, the reward value of each outcome, and the probability of each outcome occurring) we can then calculate the expected reward using the numerical expectancy method. For the behavior of "buying of lottery tickets", the probability of winning the top prize and the value of the prize are known. Therefore, we can easily calculate the value of the expectancy reward for such behavior.

We assume that $\overline{r}_i = \{r_{i1}, r_{i2}, \ldots, r_{in}\}$ is the set of reward values in one-to-one correspondence with outcomes of behavior i, and it is a mutually exclusive outcome set. A mutually exclusive outcome set is a set where one and only outcome would occur.

p_i is the correspondent probability set (there is bijection between set $\overline{r_i}$ and set p_i), $p_i = \{p_{i1}, p_{i2}, \ldots, p_{in}\}$, $\sum_{j=1}^{n} p_{ij} = 1$, then the numerical expectancy of the reward would be:

$$r_i = \sum_{j=1}^{n} p_{ij} r_{ij} \tag{2.1}$$

2.3.1.2 The Precedence Expectancy of the Reward

However, in management practice, more often than not one is unable to know the specific outcome probabilities and corresponding reward values for a certain

behavior. Under such circumstances, where the value of numerical expectancy cannot be calculated, the individual can only look towards previous instances of rewards on similar behavior by others as a way of estimating rewards. This is what we call "precedence expectancy" in this volume.

For example, with the behavior "attending graduate school", individuals cannot possibly predict with precision the employment and remuneration outcome for life after graduate school. However, they can look at individuals who have completed their graduate studies and are now in the workforce to estimate how they might fare if they were to take the graduate school route. This is what we call "precedence expectancy" in this volume.

In reality, it is more common to see individuals make use of precedence expectancy to estimate the reward on a given behavior. When selecting commodity types for trade or when selecting majors in college, it can be difficult for the actor to forecast behavioral rewards using precise data. However, they can determine the rewards from various possible choices using the precedence expectancy method.

2.3.2 Issues to Note When Using Reward Measures

2.3.2.1 Rewards and Penalties Should Be Well Grounded and Pre-established

From the perspective of acceptability and recognition, institution rules should be pre-established and made clear from the onset together with the corresponding rewards and penalties. This way, the institution is more likely to be accepted by individuals (both within the institution and outside the institution). Arbitrarily granted and meted-out rewards and penalties will only incur the displeasure of individuals and would not be very well accepted. Therefore, decisions for the granting and meting out of rewards and penalties must be grounded on a strong basis with the relevant rules created and made clear to everyone in the target group beforehand.

2.3.2.2 Measure Accuracy, with the Use of a High-Performance Observer

When we use reward measures, there is the question of when a reward should be given, when a penalty should be meted out, and how generous or severe should such measures be. This is an issue of accuracy for the reward measure.

To enhance the accuracy of the reward measure, when performing institution design, the designer must ensure that a high-performance observer is included in tandem with the reward measure (i.e., the reward-type promoter). For example, reward and penalty standards should be made clear and target behaviors should be easily observable with accuracy.

2.3.2.3 The Prevention of Reward Compensation

The term "reward compensation" refers to a situation where the manager can only control part of the reward to the actor for a given behavior. This way, the manager's control over actor behavior would be greatly diminished.

For example, in families where only the husband is working, the husband would work extra hard as he is the sole breadwinner. With families where both the husband and wife are working, especially when the wife draws a very high wage, reward measures such as rewards and penalties would not work as well on the husband in the workplace as rewards from work would not matter to him as much as if he were the sole breadwinner.

2.3.2.4 Improving the Detection of Positive Guiding Rewards: The More Direct the Reward Chain, the Better

The term "positive guiding reward" refers to a reward that guides the actor into selecting the target (desirable) behavior. Within an institution environment, the term refers to the reward for specific behaviors as mandated by institution rules.

The term "reward chain" refers to the various segments or stages that occur between the occurrence of behavior and the obtaining of behavioral rewards by the actor. If there are many intermediate or intervening segments [between behavior and reward], then the reward chain is considered long. An example of a reward chain in the case of an employee receiving a cash bonus for his hard work would be: Hard work → Work performance → Recognition by superiors → Cash bonus received.

The longer the reward chain, the higher the uncertainty of obtaining the expected reward following an action (behavior). For example, the uncertainty may hinge on whether the behavior in question was observed, or whether the bonus measure would be implemented properly, etc. The more the number of intervening segments, the lower the probability of obtaining expected rewards following an action (probability is multiplied across segments, with probability generally lower than 1 and decreasing with each intervening stage). On the other hand, the shorter the reward chain, the more likely that the actor would receive the expected rewards. The more observable the reward on the part of the actor, the stronger the guiding effect of said rewards. Conversely, the less observable the reward on the part of the actor, the weaker the guiding effect of said rewards.

For example, in the case of the use of banned food additives by food manufacturing enterprises, the relevant state agencies have put in place corresponding negative rewards for such actions. However, the reward chain in this case is long: first, the addition of such substances to food products must be observed. This typically happens with scheduled sampling and tip-offs from the public, which means that the contravention of this rule may or may not be discovered in the first place. Even if such behavior has been discovered, an enforcement team is needed to mete out the corresponding penalty to the offending manufacturer. There is no

guarantee that the enforcement team would take its work seriously in this regard. We see that with a long reward chain such as this, it is no longer a surprise as to why the behavior of using banned additives in food products by manufacturers continues to be a rampant problem.

On the other hand, when individuals make their own food at home, they do not need any observation or penalties in place, and yet people generally do not add banned substances to their food. That is because if they do so, undoubtedly they would end up harming themselves. This is a very direct kind of reward with no intermediate links or stages.

Similarly, objectively speaking if a negative guiding reward chain (i.e., a reward chain that causes undesirable behavior) cannot be completely cut off, then we can try to lengthen the reward chain in order to weaken its effects.

2.3.3 Reward Interference and Promotion in the Institution

2.3.3.1 The Three Behavioral Tendencies of the Actor

The basis for the selection of any behavior by the actor is the maximization of behavioral utility. Therefore, the actor has three behavioral tendencies: "pro-self", "pro-others", and "oppositional". Among these, "pro-self" refers to a tendency for behavior that benefits the actor himself, with economic rewards and social-standing rewards typically providing the motivation for such a tendency. "Pro-others" behavior, which benefits a specific target (individual or group) other than the actor himself, arises from attitudinal and emotional rewards in many cases. The "oppositional" behavioral tendency of the actor refers to the tendency to behave in a way that is unfavorable to a target. Attitudinal and emotional rewards are also often a cause of this tendency.

From the perspective of institution effects, behavioral rewards may lead to two different effects: institution interference and institution promotion.

2.3.3.2 Reward Interference in the Institution

Behavioral rewards can cause interference within the institution, if institution implementers choose poorly (such as when certain rewards cause the implementer to opt for a behavior unfavorable to the institution) and the observation of institution implementers is poor or when there are inadequate limits on the behavior of institution implementers (such as reward-type, resource-type or opportunity-type limits).

There are many instances where reward measures have led institution implementers to adopt undesirable behavior. For example, in institutions where individuals that are part of the target group have been allowed to become implementers of benefit distribution institutions—such as the lottery—then the "pro-self"

tendency when realized as the behavior can create a "positive reward" for the implementer. Under such circumstances, institution implementers would find it hard to "distribute fairly", leading to implementation bias within the institution.

To take another example, if the institution implementer has had a conflict with an individual under management (which typically leads to the individual being punished overly harshly or receiving less rewards) or when there is a family relationship between the implementer and an individual under management (which typically would lead to the individual getting off with light punishments and receiving comparatively more rewards), etc., then as all three behavioral tendencies on the part of the aforementioned actors exist, it may lead to implementer behavioral bias instead. In such situations, the "related-parties" policy is a method commonly used to prevent such interference.

In addition, "rent-seeking behavior" on the part of those with power, i.e., institution implementing behavior by institution implementers that "serves the self" for positive rewards turns authority that is meant for "the impartial discharge of one's duties" into a tool for self-enrichment.

Example 2.7 The manipulation of numbers by officials

Numbers maketh the official, and the official maketh the numbers. The phenomenon of falsifying numbers and of over-stating performance is a classic one in officialdom. This problem has remained widespread because although everyone knows that such behavior is "undesirable behavior", it is not in the interest of superiors to actually deal with such problems honestly, as this would in turn lead to the "deterioration" of the superior's score sheet since the performance of the superior is partly a composite of his subordinates' performances. This is effectively a behavior that presents reward interference for the superior's "punishment for subordinates' misdeeds" (though managing these subordinates was indeed the job of the superior). Hence, many superiors are not too proactive when it comes to tackling such behavior by their subordinates. This type of reward interference is one of the main reasons why we see the practice of the falsification of performance data permeate certain agencies at various levels.

Example 2.8 The "Fair Scale"

For the protection of consumer interests, a "Fair Scale" was set up at several marketplaces in a particular city. The scale, which typically was managed by a dedicated staff member, was put in place for customers who wanted to verify that they had received the amount of goods they had paid for.

However, during an inspection by the city, it was found that not only did many administrators of the "Fair Scale" fail to protect the interests of consumers, some of them had even conspired with marketplace vendors to cheat the consumer. In one instance, a consumer had gotten into an argument with a rice vendor who had shortchanged her. At this point, the "Fair Scale" administrator came along, and told the customer: "It is not like you are buying gold. Why are you being so niggardly?"

That the "Fair Scale" administrator would gang up with businesses to shortchange the customer was something the relevant agencies had not foreseen.

Commentary:

This problem arose due to the selection of undesirable behavior the institution implementer caused by reward interference. When consumers complain of being shortchanged, the administrator of the "Fair Scale" can make one of two behavioral choices: to protect the interests of the consumer or to side with the vendor.

The rewards on these two types of behavior are different. As consumers are but transient passers-by at the marketplace, the administrator would not gain very much by protecting the consumer's interests. On the other hand, the administrator meets the vendors at the marketplace on a daily basis. Therefore, the negative reward associated with offending these vendors (such as getting into trouble with a vendor) is comparatively greater. Moreover, siding with an errant vendor would not likely result in any negative reward from the customer but would instead earn brownie points with errant vendors. Who knows, perhaps the administrator would be getting freebies from these vendors later on?

Therefore, of the two behavior choices available to the administrator, the reward associated with "siding with the errant vendor" is comparatively greater. And this is how implementer bias occurs under reward interference.

2.3.3.3 Institution Promotion Using Rewards

Under certain circumstances, rewards can also be used to promote a desirable institution outcome when used appropriately.

For example, when the victims of official corruption take on the task of "observing corrupt behavior among officials", as such behaviors have a direct impact on their interests (this is the negative reward inflicted on them as a result of corruption), these "monitors" would work very hard to complete their tasks.

Example 2.9 The "Discharge in the Open" Policy

A certain city has implemented a "discharge in the open" policy with enterprises, meaning that they are required to send out all discharged through known discharge channels equipped with instruments that measure discharge flow, chemical oxygen demand (COD) levels, pH levels, etc. This way, the enterprise's discharge activities would be out in the open and the public would be able to monitor the activities of the enterprise at any time using readings from the instruments installed.

After the policy has been implemented, illegal discharge activities have more or less been eradicated in the area while the number of cases of discharges that do not meet environmental standards has also decreased significantly. In many towns and cities, apart from the professional environment control personnel sent by the state, many members of the public also took it upon themselves to become "volunteer environment control personnel". Recently, when a chemicals factory created a secret discharge channel, residents living nearby reported the enterprise with the result that the enterprise was closed for a period and fined.

Commentary:

As members of the public suffer from pollution (i.e., a negative reward) caused by enterprise discharge, they tend to be the most motivated when it comes to the monitoring of enterprise discharge activities, with some even becoming "volunteer environment control personnel". In this scenario, monitoring by members of the public is comparatively more effective. This is the promoter effect of the reward within the institution.

2.3.4 *Reward Measures Often Need to Be Paired with Bias-Prevention Measures*

The reward measure can enhance the proclivity on the part of the actor towards a certain behavior. This is why it is a widely used management measure. However, in many situations, the reward measure can promote more than one type of behavior by the actor. Therefore, the accuracy of the measure is typically not high, and thus the measure may induce the actor to opt for a certain other undesirable behavior instead. Under such circumstances, a certain bias-prevention measure (such as another reward measures or a resource measure or an opportunity measure) should be used in tandem to prevent the actor from showing behavioral bias.

Example 2.10 Case fabrication by special agents during the "White Terror" period in Taiwan[1]

During the "White Terror" period in Taiwan (following the retreat of the Kuomintang from the Chinese Mainland to the island), "any special agent who successfully caught a "criminal" would receive a large bonus and have the opportunity for a job promotion". This resulted in Kuomintang agents to regard "the arrest of 'undesirable elements' as a way of climbing the job ladder". In one example, several students in a naval school were accused by special agents of pledging to work for the Communist Party simply because of a letter that read: "We must study hard and stay on the straight and narrow, for anyone can become a useful person" (Chen 2013).

Another person who had been involved in such shenanigans claimed that some of the cases "manufactured" were "not products of the judiciary but created by 'artists'". At that time in Taiwan, intelligence officials would describe cases as being "made [up] well" or "not made [up] well".

A classic example of this practice was the case of the "713 Penghu Incident" that had rocked Taiwan at that time. During the Civil War in China, over 10,000 students of the Yantai United High School fled to safety in Penghu, Taiwan. Due to manpower shortages, the local Kuomintang branch in Penghu made the

[1]Details for this case study were excerpted from Chen (2013). Narrative has been re-written for the purposes of this volume.

conscription into the local infantry regiment of students not yet sixteen years old but otherwise declared "fit for conscription" mandatory. This resulted in bloody clashes between unwilling students and the military. The special agents then set to "work".

They first needed to find a "reason", and this "reason" was to be that there was a large Communist spy organization within the school that was sabotaging the Taiwanese conscription effort.

Then, they went about collecting "evidence". They began from the bottom rungs because individuals at the bottom yielded more easily. Then, connections were made up the food chain.

With the "evidence" collected, it was time for the piecing of the narrative. Case investigators interrogated arrested students for hours on end, then cherry-picked parts of conversations for case use like an author would do in writing a novel "based on true stories".

Five individuals were picked from the arrested students and then observed for their respective personalities. Their "roles" in the "Communist spy organization" were crafted based on such observations. This way, the one who wrote well became a "propaganda member of the CCP", and the physically strong fellow became "the one who was in command during the unrest". Hence, these students became the "deputy troop leaders of the Yantai New Democratic Youth League". Then, their confessions became the basis for the arrest of other "conspiring" students.

The narrative would only be convincing if more members of each troop were presented for the case. The five were first given the opportunity to think about who their "troop members" were. If no names were offered, investigators then presented the "deputy troop leader" with a name list.

Any talk of a large spy organization consisting of only five student leaders would not be very convincing. "Bigger fish" were needed by the case investigators. And so, the principal of the high school became a "member of the CPC Jiaodong Executive Committee" while the deputy principal became "member of the CPC Yantai Party Committee-cum-Head of the Yantai New Democratic Youth League".

Clearly, torture was used to extract "confessions" during this process. In the short span of just over a month, case investigators were able to put together a case thanks to the use of nine torture methods. Finally, the case was handed to the Security Command in Taipei, and the two principals along with the five students were sentenced to death on the charge of "conspiring to overthrow the government using illegal means".

Said the person who recalled the event: "Under torture, everyone was willing to work with the investigators' narrative and thus created identities for themselves. They then also corroborated each other's claims. With an identity made up, next came the actions. Again, they corroborated each other's narratives, forming an intricate web of relationships within the narrative. This web of relationships had an internal logic, with each part highly interdependent, [and the entire narrative] complete enough to stand on its own."

Commentary:

In Taiwan, the opportunity for Kuomintang special agents "who successfully catches a [guilty] individual would receive a large bonus and the opportunity for a job promotion" was taken as "a stepping stone towards job promotions and greater wealth". Clearly, this was a reward measure adopted by the Kuomintang at that time to motivate its special agents. However, with this reward measure, special agents had two possible behaviors to choose from: to work hard at solving cases in the honest way, or to fabricate cases. The problem was, in a situation where there were not many cases to solve, only by "fabricating cases" could an agent hope to move up the ladder. Hence, the use of just one reward measure in this case caused behavioral biases on the part of the special agents. The only way to ensure that the special agents did not abuse their powers was to implement a stringent judicial audit program in tandem. However, in reality, due to the lack of such a program many miscarriages of justice were committed during the "White Terror" period in Taiwan.

References

Chen X (ed) (2013) Just how terrible was the 'White Terror' in Taiwan? Tencent History 187, 17 July 2013. http://view.news.qq.com/zt2013/bskb/index.htm

Duan S (1998) Identifying coal thieves using computers. Xinmin Evening News, 16 Dec 1998

Gao F (2011) Preventive measures 'Tofu-dreg Projects' during the Ming dynasty. Health Herald, 30 Mar 2011

Chapter 3
Observers

3.1 Institution Components

Institution components are the basic structural components of an institution.

They act as the implementors of certain specific functions within the institution. This implementer may be a smaller institution in itself, or may comprise solely of an implementing agency or of implementing equipment.

The "measure" mentioned in the previous chapter is conceptually distinct from, and related to, the concept of the "institution component".

The measure is typically a specific means designed to regulate the behavior of individuals targeted for management. Examples include reward measures and cost measures. Institution components are usually institution units with standard functions. For instance, the behavior promoter is capable of encouraging the occurrence of a target behavior or enhance the level of effort put into the said behavior; the behavior suppressor is capable of preventing the occurrence of a target behavior or to lower the level of effort put into such behavior; and the behavior observer works by detecting the occurrence of the target behavior or the level of effort put in the target behavior, etc. The aim of the "measure" is clear. For instance, "the use of a reward measure" has the objective of enhancing the reward for desirable behavior or lowering the reward on undesirable behavior. Institution components clearly exist in reality and are hence good "standby options". For example, when an institution manager utilizes a "reward measure", he can look for a "reward-type promoter" in reality to implement this measure.

In addition, the use of "measures" is a one-off action. When we say that "so-and-so has adopted measure X", it means that the action has been over and done with. On the other hand, "institution component X" within the institution exerts an influence over time.

S. Sun, *Five Basic Institution Structures and Institutional Economics*,
DOI 10.1007/978-981-10-0344-8_3

For example, in the early days of China's reforms and opening-up in the 1980s, a double-track price institution was implemented. The result was that certain individuals with "guanxi" (advantageous relationships) profited from "goods in short supply bought at a low price". The "cancellation of the double-track price institution" was an "opportunity measure" that eliminated the opportunity for some to engage in such arbitrage once and for all. In the *gaokao* examination institution, candidates are expected to use their candidate numbers in place of their names on their answer sheets so that the exam evaluator would not have the "opportunity for cheating". This process of implementation for the "candidate number" rule constitutes an "opportunity-type suppressor" within the institution that kicks in every year during the *gaokao* period.

Based on their functional differences, institution components can be divided into four categories: outcomes, observers, promoters, and suppressors. Among these, "outcomes" are transitional components. The other three types can be further divided into small categories based on their specific characteristics and principles. For example, among promoters there are the reward-type promoters, opportunity-type promoters, resource-type promoters etc.

3.2 Observer Performance

The observer is an institution component used to ascertain the occurrence of the target behavior or the level of effort put into the target behavior. Within the institution, whether or not to impose controls on behavior rewards, resources, opportunities etc., is often determined by the information observed. Hence, observers are a fundamental type of institution component within the institution.

Observer performance chiefly comprises three factors: observation intensity, observational accuracy, and cost.

3.2.1 Observation Intensity (Sensitivity)

The observation intensity (sensitivity) of the observer refers to the probability that the target behavior would be "observed" by the observer when it occurs. If the probability is high, then the observation intensity (sensitivity) of the observer is considered to be high; when the probability is low, then the observation intensity (sensitivity) is considered to be low.

Theoretically speaking, the observation intensity of the observer ranges from 0 to 100 %. Within this range, a number of 0 % indicates that the observer is completely ineffective and unable to detect the occurrence of the target behavior while a number of 100 % indicates that the observer is completely effective and is able to detect each and every occurrence of the target behavior.

3.2.2 Observational Accuracy

The observational accuracy of the observer is mainly determined by how much the information observed has come from the target rather than noise from sources not related to the behavior of the target. For example, when using the "interview" as an observer to understand the performance of the target personnel, certain conclusions may be made based on the actual circumstances of the personnel while others may be due to self-deprecation or aggrandizement by the personnel. The latter type of information is to be regarded as noise. Clearly, the lower the noise level, the greater the accuracy and the better the observer performs.

3.2.3 Cost

Obviously, assuming that observer sensitivity and accuracy levels across two observers are the same, the lower the cost, the better the observer performance.

Observer cost is related to factors such as observer performance and the number of targeted individuals the observer is expected to monitor. Generally speaking, the better the observer performance (which means that the amount of human and material resources used would be greater), the higher the observer cost would be. Similarly, the greater the number of targeted individuals the observer is expected to monitor, the higher the observer cost would be.

3.3 Observer Types

In actual institutions, observers used in an institution can be categorized into three types based on the information collection method: direct observers, indirect observers and verification-type observers.

3.3.1 Direct Observers

Direct observers are institution components that observe the behavior of individuals targeted for management directly. Examples of direct observers include the direct quantification of work put in (such as total working hours, total number of units produced, etc.) and the supervision of work status.

Direct observers can be human agents or a combination of the human agent and equipment. For example, before the founding of the People's Republic of China, in some enterprises there was the post of "work monitor". The work monitor was a human observer whose job was to observe the "effort level of employees at work". On the other hand, traffic cameras and cameras installed at bank teller counters together

with the human personnel who would examine the footage from such cameras from time to time to determine instances of rule contravention make for direct observers that comprise a combination of the human agent and observation equipment.

The sensitivity and accuracy of direct observers are comparatively high. However, as this type of observer works by directly monitoring the target, costs are also relatively high. Moreover, the scope of observation is comparatively small and the observer can only be used in instances where information can be collected directly.

Example 3.1 Cheating and invigilation in the scholar examinations of ancient times (Ni 2012)

In the imperial examination (the scholar examinations of yore), successful candidates had the opportunity to become court officials. Therefore, the outcome of the examination was highly important to candidates. Certain candidates cheated in order to gain success at the examinations.[1]

In ancient times, candidates at imperial exams were confined to a room each by themselves for the exam proper. Therefore, one major difference from the exams of today is that there was no way candidates could cheat by "slipping notes to each other". Therefore, the most commonly-seen method of cheating was the smuggling of notes into the examination venue.

This method was known in the past as *huaijia*, or "to bear to the breast". In the *Jing Lin Notes*, Zhou Fujun of the Ming Dynasty described the many ways in which *huaijia* was performed: "[Cheating candidates] would bring in minuscule booklets of a thousand pages and measuring not more than an inch thick, or hide notes inside the hollow stems of their ink brushes or under their ink stones. Other methods include notes hidden in the sole of their straw sandals; and then some have written notes in the collars of their clothing or belts using a mud formulation, with the text appearing once dirt on the belt has been removed. This method is known as the [exam] hall back-up."

With such ingenuity displayed by cheating candidates, exam administrators over the ages paid special attention to the monitoring of exam proceedings. It was recorded in historical texts that during the Jurchen Jin Dynasty, candidates were required to take off their clothing, release their hair from hair knots, and even have their ears and noses searched as part of pre-exam checks. The purpose of having candidates undo their hair knots was to check for the presence of hidden *shengshu*, a type of cheating device.

During the Qing Dynasty, the Emperor Qianlong had soldiers search each candidate, with each individual who uncovers a case of cheating promised a monetary reward. Searches were so stringent that even the buns and biscuits brought in for sustenance by the candidates were broken or cut apart for checks. In another historical record from the time of the Emperor Qianlong, it was recorded

[1]Case study excerpted from Ni (2012). The case study has been amended slightly for the purposes of this volume.

that year, for a Jiangnan (area south of the Yangtze River) examination session, 21 candidates were found to have had huaijia in the first exam alone, with another 21 caught in the second session. As a result, the gaol located inside the examination institute was packed full. For that specific exam session, eventually a total of 68 candidates turned in blank answer scripts, 329 turned in incomplete scripts, and 276 candidates were found to have provided irrelevant answers in response to the essay question. At the second session, a total of 2800 candidates were found to have dropped out from the exam.

3.3.2 Indirect Observers

Indirect observers are institution components that observe the behavior of individuals targeted for management through an intermediary. Examples of the use of indirect observers include the use of survey questionnaires and tip-offs to understand the behavior of target individuals.

Because the observation of the behavior of target individuals is not conducted directly, the costs of indirect observers are lower. On the other hand, the scope of observation is also broader. However, the sensitivity and accuracy of this type of observer is comparatively lower. For instance, during an investigation with "targets for career fast-tracking", inaccurate self-narratives were common, and false information including "false accusations" was also found in reports.

Example 3.2 Tip-offs in ancient China (Zhou 2006)

The "tip-off" institution has had a long and illustrious history in China. Historical records have it that over 5000 years ago, in the time of the mythical Emperors Yao and Shun, pillars were erected along artery roads on which people express their opinions through carving. These pillars were hence also known as "accusation pillars".

During the Western Han period, a bamboo canister known as the *jitong* was used for tip-offs. The *jitong* is a canister with a hole at the top through which a note would be placed in but not removed. This is probably the earliest instance of a "tip-off box" in China. In the *Book of Han*, it is recorded: "Officials were asked to make *jitong* for note collection".

Under the reign of the female emperor Wu Zetian, the "tip-off box" evolved into a more sophisticated form. The device, known as the *tonggui,* was a square copper box with four openings, one on each side. The cyan opening on the east side is for the dropping of laudatory odes, while the red opening on the south side was for individuals wishing to put in an honest word with the emperor. For those wishing to right a wrong, the appeal was to be dropped into the white opening on the west side, while documents of a sensitive military nature were to be dropped in from the black opening on the north side. The *tonggui* was a key way for Wu Zetian to understand developments outside the confines of the palace walls, and she made use of device to consolidate her rule.

3.3.3 Record-Type Observers

The use of record-type observers involves the recording of behaviors of target individuals and then examining the "records" of such behavior when necessary. Various certificates (which state the qualification held by the holder) used in the modern society, official document stamps or authorized signatures (proof that the relevant processes have been completed), engraving of the maker's name on the product (for traceability in the event of product defect) etc., are all record-type observers.

Record-type observers are not limited by time and space and are of low cost. Therefore, they are used widely in modern society. However, this type of observer can be easily forged.

Example 3.3 Verification using coded wooden blocks (Tan 1998)

In July and August 1998, a massive flood occurred along the banks of the Yangtze River. The dikes had to be patrolled day and night so that developments of concern could be spotted quickly and tackled in time.

However, as time wore on, many personnel along the long stretch of the river became complacent, and signs of a perfunctory attitude displayed by patrollers began to emerge. Patrollers tended to miss out on checks on mandated locations in mushy spots, poorly-lit locations and in the forest and then claim that they had indeed performed the requisite checks. It was impossible for managers to determine the veracity of these claims. Therefore, the design of a good observer of patrolling behavior was the key to the success of the anti-flood effort.

At this point, at Sanfengsi Town in Huarong County, the locals devised a "password-check method" operated by individuals familiar with the local dikes. Under this method, a number of wooden blocks marked with numbers were placed at locations that patrollers were required to visit. This way, patrollers had no choice but to be serious about their work as they were required to locate the coded blocks and submit them to the supervising office. After this method was implemented, the effort level put into their work by patrollers rose significantly, with the result that ten instances of impending danger were uncovered. A number of these cases were specifically uncovered due to the placement of the coded blocks.

Commentary:

This is also a record-type observer, as the coded blocks serve as a record of patrollers' checking-in at certain spots, and of the level of effort put into the desirable behavior. The key to the performance of this observer was: the wooden blocks could not be forged by patrollers beforehand as they could not possibly know what the numbers marked on the respective blocks were. Only if they had seen the wooden blocks in question would they be able to produce the correct numbers at the end of the patrol. This carefully designed observer presents rather accurate observations of the effort level put in by patrollers during their patrols. From this example, we can see that just like other scientific endeavors, the task of institution design also requires innovation and creativity.

3.3.4 Response-Type Observers

The response-type observer works thusly: first it sends out a signal to the target under observation. Then, the target produces a response to this signal. The observer then detects this response to determine the specific status of the target.

Typically, examinations test for the response of exam-takers to help determine the level of knowledge or subject mastery on the part of the test-takers. Within the context of an institution, exams would be regarded as a response-type observer.

The test of an individual's response (behavior) in a specific setting in order to determine the individual's thoughts, beliefs, situation etc., effectively works on the same principle as the response-type observer.

Example 3.4 The story of the lime circle

In this story from the Song Dynasty, Squire Ma had just passed away, leaving behind two wives. Of the two, only the second wife had given birth to a son, which means that the family business would likely fall into the hands of that branch of the family.

However, the first wife then conspired with the local magistrate Zhao Lingshi to seize the child and to drive the second wife out of the family compound. The case went up all the way to the capital Kaifeng, and came under the purview of Justice Bao. In court, both women insisted that they were the mother of the child.

Justice Bao decided to take another tack. He instructed for some lime to be brought, and had a large circle on the floor of the court drawn using the lime. He then had the child stand inside the circle, with the two women on each side of the child and holding one arm of the child's each. He then declared: whoever was able to pull the child to her side would have custody of the child.

At the command "Start!" from Justice Bao, the first wife yanked at the child. On the other side, seeing how thin and vulnerable-looking her child's arms were and how the first wife was exerting plenty of strength, the second wife hesitated. The child was pulled over to the first wife's side of the circle in a flash.

Seeing that the second wife did not make an effort, Justice Bao ordered the tug-of-war to be conducted again.

The second time round, the first wife continued to tug hard at the child's arm.

The second wife did the same. However, the child then started crying in pain. She then loosened her grip on the child, and the child was pulled over to the first wife's side.

The second wife sobbed as she spoke of the reason for her refusal to pull hard at the child. Everyone present in the court was moved and convinced that the second wife was the true mother of the child.

Slamming the bench, Justice Bao ordered the first wife to speak the truth, and the case was finally solved.

Commentary:

Two women, one the real mother of the child and one the usurper were put to a test by Justice Bao. The ensuing tussle showed that the real mother cared most about her

child's safety while the usurper was most concerned with being able to pull the child to her side. The respective behaviors of the two "mothers" showed the difference in the respective relationships the two women had with the child. With the command "pull the child over to your side of the circle", Justice Bao successfully elicited the response that allowed him to tell the truth of the matter. The use of this device by Justice Bao to determine the truth of the matter also reflects the working principle of the "response-type observer".

3.3.5 Verification-Type Observers

Verification-type observers work by exploiting the inherent relationship between bits of information (this relationship can be a natural one, or one that has been set up) to observe the veracity of information received. For example, a superior may collect information on the same matter from two sources separately. If the two reports do not coincide in content, then at least one of the two reports must be false. At this time, the relevant information must be checked again. Conversely, if both reports coincide in their respective content, then the chances of both reporters being wrong are lower, which means that the probability that the information received is accurate is greater. When the key information is being entered into a database, very often the user is asked to enter the same information twice. The two entries are then compared, and only when they match is the information entered stored in the database. This approach serves to prevent data entry error.

In management practice, the numbers of certain documents are according to an "authentication institution". This also utilizes the principle of the verification-type observer.

Example 3.5 Checking the legitimacy of an identity card through a verification of its number

China has adopted a verification code institution as a way of tackling the forgery of citizen identity or ID cards. The specific method is as follows:

First, take the first 17 digits of the ID number and multiply each number by a specific factor. The factors to be used, in order from the 1st to the 17th digits, are: 7, 9, 10, 5, 8, 4, 2, 1, 6, 3, 7, 9, 10, 5, 8, 4, and 2. Then, add up the 17 numbers obtained following multiplication, and divide the sum by 11 to obtain the remainder. Finally, one can then look up what should be the 18th digit of the ID number in a substitution table using this remainder. With this method, one can obtain the 18th digit of any ID number using the first 17 digits of the number. Then, we can match the 18th digit obtained through calculations and substitution with the digit on the card for the preliminary test of the legitimacy of the ID card.

The specific substitution table for ID numbers is as follows:

Remainder	0	1	2	3	4	5	6	7	8	9	10
Substitute	1	0	X	9	8	7	6	5	4	3	2

Take for example the ID no. 34052419050104001X. Let us check to see if this number is legitimate.

First, we multiply the first 17 digits of the number by the factor and then add up the resulting numbers:

$$7 \times 3 + 9 \times 4 + 10 \times 0 + 5 \times 5 + 8 \times 2 + 4 \times 4 + 2 \times 1 + 1 \times 9 + 6 \times 0 + 3 \times 5 + 7 \times 0 + 9 \times 1 + 10 \times 0 + 5 \times 4 + 8 \times 0 + 4 \times 0 + 2 \times 1 = 171$$

Then, we calculate the remainder of 171 divided by 11 to obtain $\frac{171}{11} = 15 + \frac{6}{11}$. The remainder is 6. Then we look up the remainder substitution table for the remainder 6; the 18th digit of the number should be "6". However, the last digit of this ID number is "X". Clearly, this is a forged ID.

3.3.6 Reward-Type Observers

The reward-type observer uses the principle of "balanced separation" to create a reward response rule targeted at the individual under observation. This rule makes speaking the truth (disclosing the truth about oneself) beneficial for the individual and lying, detrimental. Under the influence of this reward rule, the target under observation would "speak the truth" and thereby meet the need for observation to obtain accurate information.

A. Michael Spence (1943–), winner of the 2001 Nobel Prize for Economics, described this type of observer as the "information transmission model".

For instance, in the labor market, there are two types of candidates: those of high capability, and those of low capability. At the same time, in the labor market information asymmetry exists in that only the job candidate knows his specific competency level while the employer does not. If the employer is unable to distinguish candidates of one type from the other, then no matter the candidates' competency levels they will be offered the same wage. Under such circumstances, the problem of "adverse selection" arises, i.e., those of lower capability levels are very willing to take up the job while those of higher caliber levels spurn the job opportunity. This is because the average wage is lower than the value that can be created by employees of higher caliber. (The average wage is the average value created by both workers of low caliber and workers of high caliber. Therefore, it is certainly lower than the wage level that the worker of higher caliber deserves.)

However, in reality, employers tend to base wage levels on the academic qualifications of the candidate. This is effectively a reward-type observer.

For example, the remuneration institution described as follows is a reward-type observer:

The cost of learning in order to obtain an educational certificate (economic costs, time costs etc., with the cost of obtaining the certificate lower for individuals of higher caliber and higher for those of lower caliber):

High-caliber: 4; low-caliber: 8

Remuneration benefits for those with qualifications and those without (income, standing):

With certificate, 10; without certificate: 5
Pure benefits owing to individuals under various circumstances:
Benefits to high-caliber individual from education: 10 – 4 = 6*
Benefits to high-caliber individual without education: 5 – 0 = 5*
Benefits to low-caliber individual from education: 10 – 8 = 2*
Benefits to low-caliber individual without education: 5 – 0 = 5*

We can hence see that under this remuneration institution, individuals of high caliber who have opted for getting academic qualifications would enjoy the greatest benefit while individuals of low caliber and who have not opted for getting academic qualifications would also enjoy the greatest benefit possible. This creates of a balanced separation of the two groups of individuals and makes for an effective observer of an otherwise difficult-to-detect piece of information, the job candidate's competency level.

3.4 Measure Selection and Matters to Note

In structural terms, there are two key elements to the observer: the first is the measure of behavior detected, and the second is the agency or equipment that implements such observation.

The term "measure" here is an indicator of the behavior detected by the observer. During institution design, the measure is selected by the designer beforehand.

Under certain circumstances, the observational effect of the observer is chiefly determined by the behavior measure. Therefore, the selection of a behavior observer is the key.

3.4.1 The Observability of the Measure

The observability of the measure is a basic requirement. If observability is poor, then the measure would have completely lost its significance. Therefore, in practice within a management institution, often imperfect measures would be used over a period of time as they offer comparatively good observation performance with no other alternatives in this regard.

For instance, an academic's publication count is a commonly used measure in the field of research. This indicator has several flaws, such as the tendency for the researcher to focus on "short-term behavior" under its influence, thereby churning out papers of low quality to the detriment of the quality of his or her long-term research. However, "publication count" can be easily measured and the costs of doing so are relatively low (far lower than, say, conducting a "quality assessment" on each and every paper produced by the researcher). Therefore, this "measure" has remained in use despite the existence of numerous reservations.

The observability of the behavior process and behavior outcome varies across different behaviors. Where certain behavior processes are not very observable, behavior outcomes may be better observed. At this time, utilizing "behavior outcome" as measure—such as in the case of manual labor—would obtain information with ease (such as the area of land tilled) although observational costs would be comparatively high. Therefore, the outcome measure for manual labor is often in the form of "units". On the other hand, for mental laborers such as public servants, labor outcomes may not be very observable. As such, "attendance" measures such as "the number of hours worked" are used.

3.4.2 Measure Accuracy

The accuracy of the measure is how perfectly it reflects the target behavior. If there are other situations apart from target behavior that may impact or change the measure, then the accuracy of the measure is poor.

For instance, many local governments use "economic benefits for the enterprise" as an indicator for the performance of the enterprise manager. However, in truth, the economic benefit to the enterprise is related both to the effort level of the enterprise manager as well as the market environment that the enterprise is in. Therefore, when the level of economic benefits to the enterprise rises dramatically, it may be a result of the hard work of the enterprise manager or the result of a strong market environment. Hence, to simply use "economic benefits to the enterprise" as the "measure" of evaluation for the effort level of the enterprise manager would be a poor fit.

In fact, using both "economic benefits to the enterprise" and "market environment" as indicators of the effort level of the enterprise manager would make for a more holistic measure.

3.4.3 Measure Rigidity

Measure rigidity refers to how resistant or unyielding the measure is to impact by non-target behaviors. If the measure can be impacted easily by non-target behaviors, then measure rigidity is low as it can be changed easily by human manipulation.

For example, in research management, if we do not look at the quality of academic output and merely use "publication count" as a measure of "level of effort put into research behavior", this measure is not very resistant as some individuals are able to pad their publication count by various means, such as seeking the help of friends or even hiring ghost writers. In this example, "research behavior" is the behavior targeted for observation, while "seeking the help of friends or hiring ghost writers" are target behaviors that need not be encouraged but which have an impact on the measure outcome. Hence, with this measure in place, when the number of publications put out by an individual increases exponentially, the manager would find it difficult to determine if this is the result of hard work or achieved through undesirable behavior.

A better measure to use would be "count of papers published in authoritative journals", because typically the outputs of ghost writers do not meet the standards of such journals. At the same time, such journals pay more attention to the maintenance of their reputation and are more stringent in their management. Thus, aspiring authors would have a harder time to "[seek] the help of friends" in such cases. Under such circumstances, the "count of papers published in authoritative journals" would be determined chiefly by the research capabilities of the individual concerned. Therefore, this is a more resistant measure.

Example 3.6 The buying of patents by college graduates seeking household registration in a certain city

A certain city has imposed a "*hukou* [household registration] points" institution for college graduates wishing to settle in the city. Three types of points can be earned towards eligibility: "school points" (determined by which tier the applicant's alma mater is classified in), "qualification points" (based on whether the applicant's highest educational qualification is a bachelor's degree, a master's degree, or a PhD) and "patent points" (based on patents held by the applicant, such as patents for product external design and effect models and inventions). Applicants who meet the city's total point cut-off are then eligible for household registration in the city.

Among these categories, the easiest for an applicant to manipulate would be the category of "patent points". In practice, patent-granting agencies perform perfunctory checks on applications for patents for product external design and effect models, and as long as the application shows a change from the original product it is generally guaranteed of a "patent". For instance, an application that proposes a drinking glass that is a "hollow rectangular column" or a "hollow three-sided column" instead of the traditional cylinder shape would succeed in obtaining a "patent" even if the significance of such innovation is unclear. We can see how "patent points" present a point of entry for individuals less earnest.

Indeed, certain companies have started businesses that apply for patents on the behalf of college graduates for the purpose of "*hukou* points", with the pricing of such services "rather reasonable" at RMB288 per patent as of 2011. Clearly, "patent points" can be obtained fairly easily.

3.4.4 Being Cautious in Weighting: Avoiding Interference

Currently, it has become a trend to use "integrated measures" comprising various indicators. For example, when ranking colleges and universities in China, certain units have created an "integrated point institution" with indicators such as "student population size", "publication count", "patent count", "science & tech award count", "teaching award count" "building area", "land area", "faculty size", are totaled for a composite point for sorting purposes. Institutions that appear at the top of the list with such measures may not be excellent either in research or teaching, but have achieved their position thanks to "so-so performance for all indicators". Objectively speaking, this so-called "ranking" signifies nothing.

Some institutions of higher education and research even add indicators for different areas such as from "research output" and "educational background" together for the purposes of human resource management, integrating research output indicators with whether the individual in question has graduated from a "Project 985" school or an overseas institution. We all know that "research output" refers to the output of the talent, and is the only standard by which we can measure the talent's capabilities. The factor "if the individual has graduated from a 'Project 985' school or an overseas institution", on the other hand, is related to the individual's background. This type of "weighting" has, in practice, allowed individuals to compensate for lower research output with their stronger backgrounds and obtain job promotions alongside more capable individuals. Apart from rendering the talent evaluation institution highly unfair, this practice has also reinforced prevailing attitudes about an individual's background. It would appear that some administrators at these institutions, even those holding PhDs, are not capable of meritocratic practice.

3.4.5 Preventing Biases from the Use of One-Sided or Unfair Indicators

On one hand, we should not abuse the "weighting" so as to prevent the dilution of "key indicators". On the other hand, we must also pay attention to biases in behavior targeted at individuals under management due to inherent biases in the indicators used.

For example, with entities such as "publishing houses" and "the media" that have needs for economic benefit yet are required to practice social responsibility, if we are to use simply "economic benefit" as a measure of the competency of their managers and administrators, this would lead to the abandonment of the entity's social responsibility as those being evaluated seek to maximize economic benefit. An example of such action would be the publishing of material promoting superstition or even pornographic material in order to boost sales numbers.

For such a situation, generally speaking the "weighting" is still not the solution to tackle such behavioral biases because as long as the "economic benefit" indicator is prominent, it can still lead to a high composite score and compensate for any loss in the "social responsibility" category.

The correct way to prevent behavioral bias should be the adoption of "threshold institution + points institution" method, i.e., applying a "threshold institution" (also known as the "veto institution") on certain lines with social impact. With reference to the example just used, we can impose the threshold institution with regard to the "social impact" indicator, which means the publishing house may not put out publications injurious to society. If the publishing house is caught doing so, it shall also receive a "0" in the "economic benefits" category. Only when there are no issues with the "social impact" indicator would the enterprise then be scored on the "economic benefits" indicator.

3.5 Information Bias on the Part of Information Collectors

For observers used within an institution, there must be an information collector who is responsible for the collecting or provision of information needed. Different information collectors may have different biases leading to the loss of certain information. We can also categorize observers by the type of information collectors used.

3.5.1 Observers that Make Use of Dedicated Information Collectors

Dedicated information collectors are often used in some observers. For instance, during the "people's communes" era, there were "time clerks" who were responsible for recording the attendance of commune members while "weighers" were responsible for weighing trucks filled with coal when mine output was being shipped out of the mine premises. These are both examples of dedicated information collectors.

A supervisory institution must be established to ensure that dedicated information collectors would "collect information impartially". If the supervisory institution is not in place or implemented poorly, issues to do with corruption such as "rent-seeking behavior" may arise with dedicated information collectors. Of course, the use of dedicated equipment to serve as "automatic information collectors" is another effective way of dealing with corruption amongst dedicated information collectors.

3.5.2 *Reporting-Type Observers*

The information collected by the reporting-target observer is provided by the target individual. For example, some company employees clock in by having their fingerprints scanned. This is a "reporting-type" observer. Observers that make use of "tests" and "interviews" to collect information are also regarded as "reporting-type" observers.

As the information obtained by the observer impacts on the interests of the reporting individual, the information obtained by this type of observer often presents as bias as reporting individuals seek to report information that is beneficial to themselves. To prevent the occurrence of such situations, we should try as much as possible to turn such reporting into objective data. For instance, the use of "fingerprint scanning" in attendance-taking is a kind of objective record that the individual under observation would find difficult to forge or manipulate.

Example 3.7 GDP numbers from local authorities versus GDP numbers from the central government (Ye 2013)

Numbers maketh the official, and the official maketh the numbers. This is a classic situation where public officials make up or fudge numbers [in order to appear more capable or to be better able to answer to their superiors]. This phenomenon can be very difficult to tackle.

In early 2013, 2012 GDP (Gross Domestic Product) as reported by various provinces totaled RMB57.69 trillion. This was RMB5.76 trillion higher than the national GDP figure of RMB51.93 trillion measured by the National Bureau of Statistics, a sum equivalent to the GDP output of Guangdong Province.

Since GDP began to be calculated by both the state and local authorities in 1985, the total of GDP numbers provided by local authorities has always exceeded state numbers, with the situation becoming worse by the year. When totaled, the GDP numbers claimed by respective provinces exceeded centrally-tallied GDP numbers thusly: RMB2.68 trillion in 2009, RMB3.2 trillion in 2010, RMB4.6 trillion in 2011, and a whopping RMB5.76 trillion in 2012.

However, local authorities do not always inflate local GDP numbers. How they choose to report their GDP numbers depends on the specific "purpose" of the statistic. For example, some local authorities would inflate their numbers for the purposes of performance evaluation, and then deflate GDP numbers when it is time to seek subsidies from higher-level authorities. One amusing example would be the case of a particular county that has been named both "Top 100 County for Economic Strength" and "Economically Disadvantaged County".

Commentary:

In this example, we see the clear problems associated with report-type observers. As GDP numbers have a direct impact on the local authority's "score sheet" for the year, aggrandizement during reporting means that there would be a bias in the information provided. Of course, the act of diminishing local GDP numbers when

seeking subsidies from higher authorities also presents information bias. The root of these information biases lies in the significance of the information to the target under observation. This is why the objectivity of the "reporting" by such targets has been undermined.

Example 3.8 The problem of case logging

For a long time, the case and resolution rates for criminal cases have been used as key indicators of work performance on the part local law and order and public security agencies. Hence, low case rates and high resolution rates are desirable targets for these public security agencies.

To ensure a low case rate and high resolution rates, in some locations, public security agencies have chosen not to officially log cases so long as they are not headline-grabbing or major cases. As a result, the use of "case rate" as a work performance indicator for local public security agencies led to the worsening of actual law and order situations in some places.

3.5.3 Observers that Make Use of Objective Information

Observers that make use of objective information typically do not present the problem of information bias. For example, observers of spending behavior on the part of members of the public depend on the collection and organization of various kinds of transaction data. Therefore, there is no information loss through bias here. Currently, in many places the collection of data by "third parties" is being advocated as the "third party" does not have an interest in the content of information collected. This is also a type of observer that makes use of objective information.

During the process of institution design, the use of observers that make sure of objective information should undoubtedly be prioritized. However, as a complete set of objective information (indicator data etc.) can be difficult to collect in many situations, the use of this type of observer should be made after careful consideration.

References

Ni F (2012) How the ancients prevented cheating in examinations. Beijing Evening News, 8 June 2012

Tan J (1998) Native invention shows power against floods. Xinmin Evening News, 16 Aug 1998

Ye Z (2013) When will the GDP fight between the central and local governments end? China Youth Daily, 4 Feb 2013

Zhou Y (2006) The whistleblowing institution in China". Qingdaonews.com. http://www.qingdaonews.com/gb/content/2006-08/22/content_7499149.htm. 22 Aug 2006

Chapter 4
Promoters

Behavior promoters, or simply "promoters" here, are institution components that are used to induce the occurrence of target behaviors or to enhance the effort level for such behaviors.

The promoter comprises two key elements: first, the promoter response rules (i.e., the functional relationship between input and output) and second, the agency or equipment that implements such rules. In practice, the promoter is often paired with an observer and typically integrated into a single mechanism to minimize costs and raise efficiency for the implementing agency. Therefore, in practice we often see institution components that comprise the integration of observer and promoter, such as with "behavior measurement + promoter response rules + observation and promoter implementer". For instance, "employee work performance indicator + rules on employee performance rewards + agency in charge of tabulating employee performance statistics and operating reward mechanism" are all institution components in an employee management institution.

As measures that can be taken to influence the behavior of the individual include reward measures, resource measures, opportunity measures, and behavioral cost measures, behavior promoters can hence be categorized as reward-type promoters, resource-type promoters and opportunity-type promoters.

Generally speaking, behaviors by individuals targeted for management can be divided into two categories: "desirable behavior" and "undesirable behavior". Institution designers use promoters to induce "desirable behavior". Therefore, in this chapter, when we talk about "effort level in behavior", more accurately speaking we are referring to "the level of effort by the individual being managed put into desirable behavior" unless otherwise specified. Similarly, we also use the term "behavior measurement" to stand for "the measurement of the level of effort by the individual being managed put into desirable behavior".

S. Sun, *Five Basic Institution Structures and Institutional Economics*,
DOI 10.1007/978-981-10-0344-8_4

4.1 Reward-Type Promoters

Positive utility is one of the prerequisites for the occurrence of the desired behavior, while the basis of such positive utility lies in the positive reward that such behavior brings. Reward-type promoters work by providing positive rewards on desirable behavior, with the link between positive rewards and the behavior in question made clear beforehand. This way, individuals targeted for management will expect the positive rewards prior to committing such behavior and may be encouraged to opt for such a behavior as a result.

4.1.1 Fixed-Reward Promoters

With a fixed-reward promoter, as long as the individual "accepts the terms of the contract" (e.g., becoming a member of a certain work unit) he will become an individual managed within the institution and will henceforth be able to receive a certain expected reward. In other words, the size of the reward received by the individual is only related to whether or not he "accepts the terms of the contract" and not with the amount of effort put into the requisite behavior.

4.1.1.1 Fixed Wages

Rules that mandate fixed wage levels plus agencies that implement this rule together make for a fixed-wage promoter. The function of fixed wages is to attract individuals into joining a certain organization. For instance, prior to the opening up of China's economy, people were particularly keen on joining "state-owned enterprises" as "formal employees" as there were a number of "fixed benefits" to be earned by doing so. However, these individuals would have no interest in working hard after joining these enterprises as there was no link between effort and wage levels.

4.1.1.2 Promotion by Seniority

In some work units, job promotion is conducted based on the workplace seniority of the worker concerned. This is actually also a kind of fixed-reward promoter, only that the key reward here is social standing (if "rank" is linked with income level, then the reward would comprise of both higher social standing and greater economic benefits).

Similarly, the purpose of "automatic promotion by seniority" is to encourage workers to be loyal to the work unit, as the longer they work in the unit, the more benefits they will enjoy.

4.1.2 Variable-Reward Promoters

The variable-reward promoter is a promoter that links "level of reward" with "level of effort put into the requisite behavior". The defining of this type of promoter is that the greater the effort put in, the higher the reward the individual can expect to receive. Therefore, before deciding to give out such rewards, the institution designer must ensure that individual effort levels can be accurately observed. Therefore, this type of promoter must be used in tandem with an observer, i.e., actual reward levels should be based on the effort levels observed by the observer and a certain mapped relationship (that is, a function that maps the effort level "indicator" to reward "value", with different promoters possibly utilizing differing mapping relationships).

Generally speaking, variable-reward promoters can be categorized into the following types.

4.1.2.1 Ladder Rewards

A ladder rewards institution divides behavior effort levels into a number of grades with a specific reward level designated for each grade.

For example, the ladder-reward structure for sales personnel at a certain enterprise is as follows.

Each person has a monthly sales quota of 1000 product units, and stands to receive a reward of an additional 1000 yuan for exceeding the quota by 100 units, 3000 yuan for exceeding the quota by 200 units, and 5000 yuan for exceeding the quota by 300 units (Fig. 4.1).

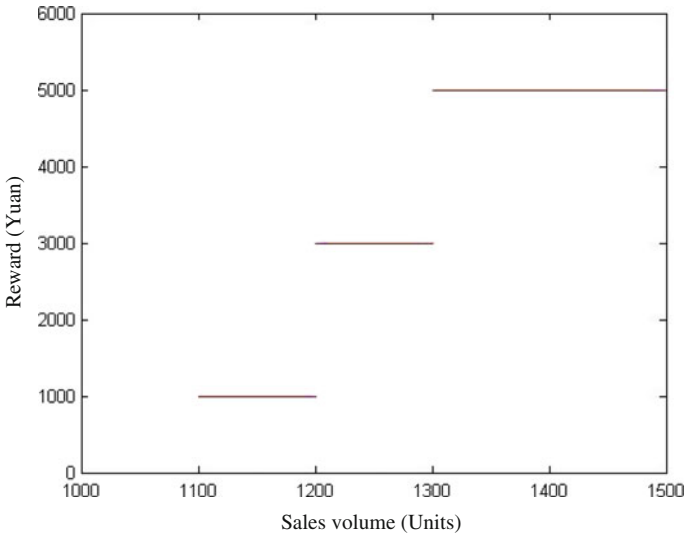

Fig. 4.1 Ladder rewards

In the history of management, the Taylor Differential Piece Work Institution invented by Frederick Winslow Taylor (1856–1915) is essentially a type of ladder-bonus institution (Fig. 4.2).
Assume that each worker has a daily production quota of 100 units.

If the worker were to produce 100–119 units a day, he shall be paid 1$ per unit produced. Hence, the worker's daily wage shall be between 100 and 119$.
If the worker were to produce 80–99 units a day, he shall be paid ninety cents per unit produced. Hence, the worker's daily wage shall be between 72 and 89$.
If the worker were to produce more than 120 units a day, he shall be paid 1$ and twenty cents per unit produced. Hence, the worker's daily wage shall be above 144$.

To use a mathematical model to describe a general ladder rewards institution, let us assume the effort level to be e_k, with e_k the increase sequence with k as the subscript. If $k < m$, then $e_k < e_m$. Set reward or reward calculation standard as r_k, with r_k the increasing sequence with k as the subscript. So, if $k < m$, then $r_k < r_m$. The expression of the mathematical model for a ladder rewards institution $r = f(e)$ is:

$$r = \begin{cases} r_1, & \text{when } e_1 \le e \le e_2 \\ r_2, & \text{when } e_2 \le e \le e_3 \\ \cdots & \\ r_3, & \text{when } e_n \le e \le e_{n+1} \end{cases} \tag{4.1}$$

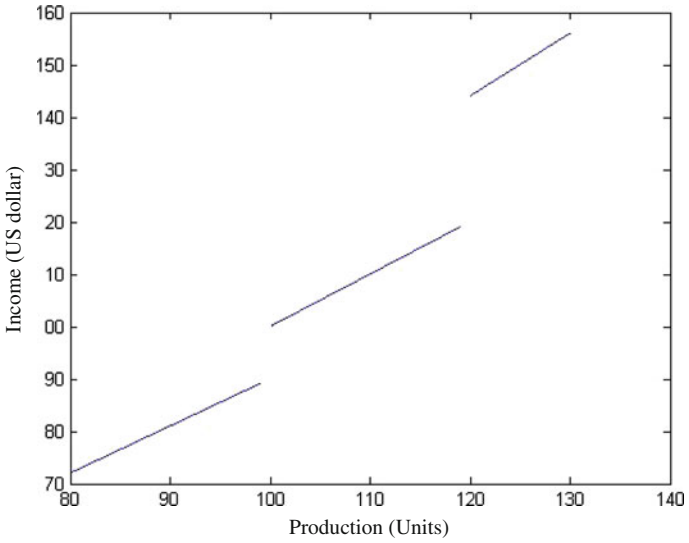

Fig. 4.2 Taylor's rewards institution

Ladder reward promoters have the advantage of being simple and straightforward, and are hence used widely.

A flaw with this type of promoter lies in the existence of a "failure interval" for effort level. When an individual's effort level qualifies the individual for a certain reward grade, any greater effort beyond this grade but not meeting the requirements of the next higher level will not result in increased rewards. Therefore, if the individual being managed feels that he is not capable enough and would not be able to meet the requirements of the next higher grade, very often he would give up. This phenomenon is particularly salient when the difference from grade to grade is significant.

4.1.2.2 Rewards Based on Sorting or Ranking

Rewards based on sorting or ranking, also known as the "championship mechanism", occur when the individuals being managed are ranked in descending order of the respective effort levels observed, with rewards for these individuals also determined by ranking such that there is a direct relationship between the two (although the relationship may not be one-to-one). For instance, the "individual with the highest effort level" would be given a "first-class award", while the "individual with the second-highest effort level" would be given a second-class award etc. In addition, granting the employee with the best performance of a job promotion or a commendation is also in fact a kind of reward based on sorting.

One advantage of rewards based on sorting is that it "filters out influences apart from effort levels", thereby enhancing the accuracy of rewards. That is, when there is a link between the respective effort levels of each individual under management or when there is a relationship between effort levels and the environment, this type of rewards institution would be more accurate that the ladder bonus institution. For instance, assuming that work performance to be the form of behavior measurement adopted, if the performance of each individual is not merely linked to individual effort but also the performance of other individuals, or when individual performance is impacted by environmental factors apart from effort (for example, the performance of enterprise managers not only depends on individual effort level but also the broader industry climate), the rewards by sorting method would be preferable.

The advantage of determining reward levels based on ranking is that oftentimes it can help to cut down on management costs. When determining individual rewards by sorting, very often only a minority of individuals, i.e., those who put in the most effort, have to be rewarded. This approach can stimulate competition between everyone in the group and spur higher effort levels across the board for significant results.

However, the rewards by sorting approach also present certain flaws. For instance, it may lead to an unhappy atmosphere within the group as the practice of ranking effectively pits individuals against one another in a zero-sum game.

In addition, when using work performance as behavior measurement indicator (and this is a very common occurrence), if capability levels vary wildly between individuals, those of lower capability levels may simply give up altogether. If the gap in capability is too great, then individuals of lesser capability would not be able to surpass those of stronger capability no matter how hard they try.

4.1.2.3 Commission-Type Rewards

Commission-type bonus institutions use a portion of the effect of desirable behavior (usually economic performance) as a reward for the individual who has turned in the performance.

Commission-type bonus institutions can be further categorized into linear and non-linear bonus institutions. When the commission is a given proportion of total performance, it is a linear institution. When the commission is determined based on grades of performance, it is a non-linear institution.

Example 4.1 Performance-based commission institution

Sales personnel at Company A are subject to a commission institution for their wages, receiving 1 % of total sales. In other words, the sales personnel stands to receive RMB100 in reward for every RMB10,000 worth of products sold. This is a kind of linear commission-based reward (Fig. 4.3).

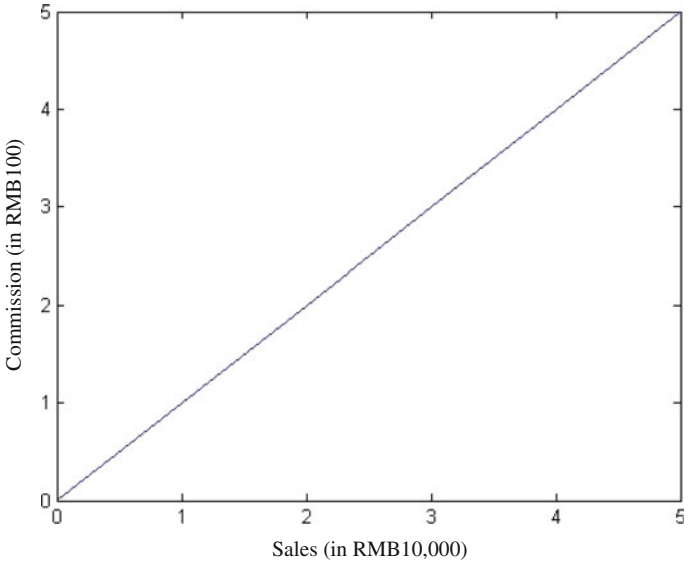

Fig. 4.3 Linear commission-based reward

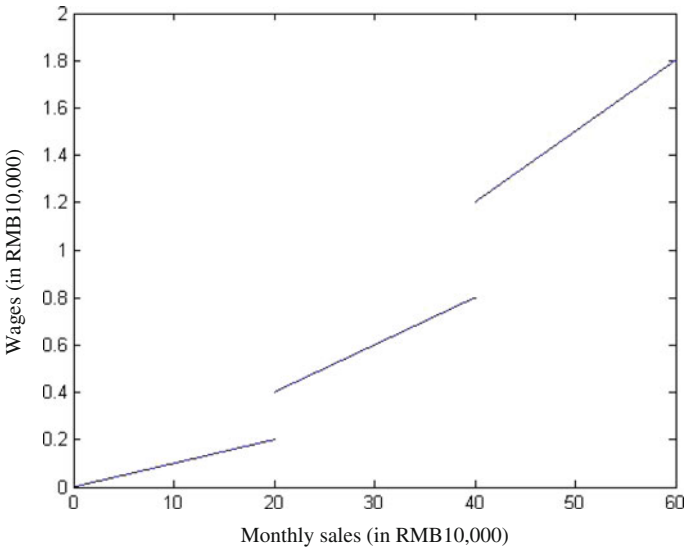

Fig. 4.4 Piecewise linear commission-based reward

Example 4.2 Performance-based commission institution based on the ladder reward structure

Sales personnel at Company B are subject to a tiered commission institution for their wages. The commission levels are: 1 % for monthly sales up to RMB200,000; 2 % for monthly sales of between RMB200,000 and RMB400,000, and 3 % for monthly sales greater than RMB400,000. This is a kind of piecewise linear commission-based reward (Fig. 4.4).

4.1.2.4 Options

Options are a kind of "reward-type promoter" aimed at managing the "long-term behavior" of enterprise managers created by enterprises in 1990s America. This type of institution component was chiefly designed to tackle the problem of "short-term behavior" by entrepreneurs.

Prior to the 1990s, the key promoters used to encourage positive "management behavior" on the part of enterprise managers were "ladder rewards" and "commission-type rewards". The advantage to these promoters is that they tend to present a stronger incentive for target behavior. However, these promoters also tended to encourage "short-term behavior" wherein managers focus on performance within the evaluation period (typically a year, and this is a kind of behavior measurement conducted for the purposes of determining rewards) to the detriment of the long-term prospects of the enterprise. Examples of such behavior include the reduction or elimination of production equipment maintenance in exchange for

higher production levels, with the net result being equipment would often have to be written off ahead of time due to wear and tear. Similarly, certain mining enterprises chose to focus solely on rich and easily accessible deposits rather than to expend time and resources on more challenging resource deposits. This type of "short-term behavior" has occurred both in China and internationally with the problem rather severe at one point.

To tackle this sort of "short-term behavior", some enterprises in the U.S. introduced the use of "options". The term "options" refers to the company's stock that is sold at a predetermined price (usually lower than the prevailing market price) to business managers. However, these business managers may not sell their holdings immediately and must wait for a predetermined period of time (usually three years or longer) to be over before they can sell these shares. This way, if enterprise managers were to focus solely on "short-term behavior" to the detriment of the long-term development of the enterprise, the share value of the enterprise would decrease, leading to a loss on the part of said enterprise managers as their options also decrease in value. Conversely, if enterprise managers are to work hard in ensuring the positive long-term development of the enterprise, they would ensure the value of their options as enterprise share value increases.

Example 4.3 Option types (Ye 2001)

There are three main types of options:

The first is the ESO or Executive Stock Option. It is given to members of the company's senior management team, and allows the recipient to purchase the company's shares at a fixed exercise price, with said shares only saleable after a stipulated period of time. The difference between the exercise price and the stock price constitutes option income.

The second type is the PS or Performance Shares. These shares are given in a certain quantity to members of the senior management team after they have met certain targets and have opted to stay with the company. PS-holders may not sell these shares within a stipulated period of time, and if they choose to leave the company or are dismissed from the company for reasons of poor performance etc., during this period, these shares would be forfeited by the company.

The third type is the SAR or Stock Appreciation Rights. These allow the holder to receive option spread income in cash terms with no need for options to be exercised. Hence, SAR is also known as "cash appreciation rights".

4.1.3 Mixed-Type Promoters

The various types of promoters described above all aim to influence behavior through a single mechanism. These various promoters bring with them certain advantages as well as disadvantages. In practice, very often promoters of various types are used in tandem for greater synergy in a mixed-type promoter.

4.1.3.1 Homogeneous Mixed-Rewards Promoter

Homogeneous mixed-rewards promoters refer to mixed promoters where the various reward-type promoters that comprise the promoter provide the same type of rewards for desirable behavior, such as when all promoters provide economic rewards or social-standing rewards.

Currently, many enterprises employ a mixed-type promoter that consists of "fixed wages + bonuses + options".

Among these promoters, fixed wages are a no-risk reward, and are used mainly to attract talent to an enterprise. As long as the hired signs on the dotted line, he is guaranteed fixed wages at the very minimum.

Bonus rewards are determined by the employee's short-term performance.

Option rewards, on the other hand, are determined by the employee's long-term performance which cannot be foreseen at the current moment but will materialize in the future. The objective of options is to prevent enterprise managers of various levels to focus solely on short-term performance to the detriment of long-term performance.

Example 4.4 Changes in remuneration structure for management personnel employed in overseas enterprises (Ye 2001)

In the 1970s, enterprise managers were paid chiefly in wages and bonuses.

After the 1980s, "manager stock options" emerged and made up 20–30 % of total remuneration. However, in recent years the value of options granted to enterprise managers has exceeded the total value of fixed wages and bonuses paid out. In a survey, *Forbes* magazine found that among enterprise managers, the median value of options granted rose 178 times from USD4947 in 1985 to USD880,000 in 1997 while the median value of fixed wages and bonuses paid out rose less than 100 % from USD730,000 to USD1.22 million in the same period. Statistics from *Businessweek* magazine show that among the top ten highest-paid CEOs in America in 1997, options income made up over 96 % of total income.

4.1.3.2 Heterogeneous Mixed-Rewards Promoter

Heterogeneous mixed-rewards promoters refer to mixed promoters where the various reward-type promoters that comprise the promoter provide different types of rewards for desirable behavior, such as with a combination that provides both economic and social-standing rewards.

For instance, the work performance of a public servant can sometimes be hard to observe. Therefore, the reward structure for public servants is commonly "fixed income (no relation to effort level, and only determined by whether the individual has joined the work unit) + job promotion (a function of effort level). This approach would attract talent to the work unit and also ensure that the individual would not lose his motivation to work after joining the unit. In the reward structure for public servants, "fixed income" is an economic reward while "job promotion" is a reward in terms of social standing. Therefore, this is a heterogeneous mixed-rewards promoter.

4.1.4 Reward-Type Promoter with Output Function

In the theoretical analysis of institution structure later in this volume (Chaps. 9 and 10), we would need an abstract but strictly regulated reward-type promoter. This reward-type promoter is best expressed in function form.

The author of this volume has designed a promoter function where the reward increases in tandem with effort level but with diminishing marginal rewards. In Chaps. 9 and 10, we shall see that this type of abstract output-function promoter can play an important role in the analysis of how institution structure impacts production efficiency.

Set the effort level of the individual i as e_i, individual i reward (or output) as r_i and the relationship function of its effort level e_i and r_i is:

$$r_i = r_0 - \frac{r_0}{1 + e_i}$$

In the equation, $r_0 \geq 0, e_i \geq 0$. r_0 is the maximum reward value, such as maximum production capacity in the case of production, the highest catch possible for a fishing boat, the maximum agricultural output possible for a specific plot of land, the point of saturation for a market etc. Therefore, the output equation has a certain basis in reality.

The reward function increases in tandem with effort level and then tapers off towards the limit. The function curve can be seen in Fig. 4.5. This is similar to the idea of diminishing marginal rewards with increasing effort level found in economics.

4.1.5 Combining Reward-Type Promoters and Observers

The principle of the reward-type promoter is: a certain reward is granted to the actor (within the institution, hence the actor under management) based on observed behavior. The response rules of the reward-type promoter targeted at the actor's

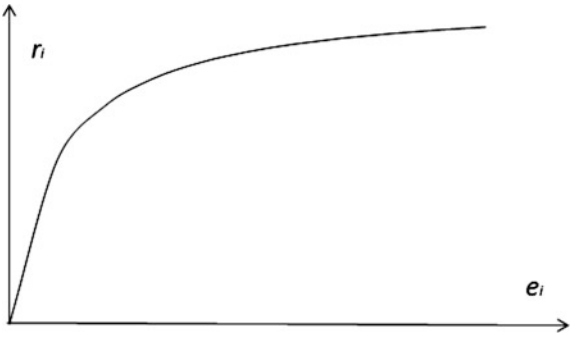

Fig. 4.5 The reward and effort level curve

behavior (i.e., the mapping between the behavior observation set and expected reward set) are public knowledge, and the actor adjusts his behavior based on his knowledge of the reward-type promoter in order to maximize the utility of his behavior.

Thus, the accuracy of behavior observation within the institution becomes a key factor in the performance of the reward-type promoter. Therefore, when utilizing reward-type promoters, specific attention must be paid to the selection of observers with strong performance to be used in tandem.

Example 4.5 The "reward-in-place-of-subsidy" policy against atmospheric pollution (Ma and Weihui 2014)

On 12 February 2014, the State Council held an executive meeting to discuss measures to tackle atmospheric pollution issues, such as the smog. At the meeting, it was decided that RMB10 billion be disbursed as part of a "reward-in-place-of-subsidy" scheme against atmospheric pollution in focused areas. The previous subsidy scheme, under which subsidies were given out based on anti-pollution plans and feasibility reports etc. drafted up, did not perform as expected. Under the "reward-in-place-of-subsidy" scheme, rewards are only given out when anti-pollution performance has been measured and deemed to have reached certain standards.

Commentary:

Originally, the state had sought to encourage local governments to tackle the atmospheric pollution through the subsidy scheme. Under this scheme, the state provided subsidies based on anti-pollution plans and feasibility reports drafted. This led to a keen interest on the part of local governments only in "winning subsidies", and these local governments would pay much attention to the drafting of such documents to the point of subsidy receipt. As to whether these local governments have put in effort in fighting pollution subsequently, however, cannot be observed effectively. Moreover, as subsidies were given out prior to "effective action against pollution", if the local government were to proceed with "action against atmospheric pollution" at this juncture, it would entail a large expense. This led to insufficient motivation on the part of certain governments in the task of fighting atmospheric pollution. Hence, the "subsidy" policy was ineffective in inducing local governments to tackle atmospheric pollution in a serious manner.

However, the "reward-in-place-of-subsidy" scheme was different in that only after evidence of effective anti-pollution measures would local governments be eligible for "reward". The effectiveness of measures taken can also be measured by the state effectively using various instruments. Here, we see that the increased accuracy of observation makes for better performance with the reward-type promoter compared to the original subsidy-type promoter.

4.1.6 The Broad Application of Reward-Type Promoters

In the practice of institution design, the reward-type promoter is the most commonly used promoter. Rules on wage, bonuses, workplace hierarchies, employee benefits and others designed to improve the work environment make for reward-type promoters together with the agencies that implement such rules.

Example 4.6 The patent institution as a reward-type promoter

Abraham Lincoln once observed that patent laws "secured to the inventor, for a limited time, the exclusive use of his invention; and thereby added the fuel of interest to the fire of genius, in the discovery and production of new and useful things." The patent institution is a reward-type promoter designed to encourage technological innovation. The economically developed Venetian Republic was the first to implement such an institution, with patent laws put in place as early as 1474. Currently, the patent institution is in broad use in many countries around the world.

The patent, which is a legal concept, essentially grants the innovator exclusive rights to the invention. The patent application is required to disclose the innovation to the public so that society at large would be kept up-to-date on innovations in the field in question. Innovations that have been patented may not be used by others free of charge, and if a patented innovation were to be used by a third party without the approval of the patent holder, the infringing party would be regarded as "in breach of intellectual property rights" and subject to penalties under state intellectual property laws. Therefore, typically it is necessary to make a payment to patent holders in order to enjoy the right of utilizing the innovation in question. This is what is called "patent purchase".

4.1.7 Comparing Rewards—Exchange Utility

4.1.7.1 Comparing Rewards

Very often, in the course of institution design there is a need to measure the impact of various rewards. If all the rewards associated with a certain behavior are of the same type, we can then directly compare the value of each reward. For instance, economic rewards can be directly compared in terms of monetary value (such as in RMB or the USD). The reward in an income of RMB20,000 is greater than the reward in an income of RMB10,000, and this is the same no matter who the recipient is.

However, we cannot directly compare rewards of different types, such as directly comparing monetary income with reputation gain. At this point, we must first estimate and then compare the utility of each reward.

4.1.7.2 The Utility of Rewards

In this volume, the "utility of the reward" refers to the actual value brought to the recipient by the reward or "the degree to which [the reward] has satisfied [the recipient's] needs".

Strictly speaking, rewards of different types cannot be compared in scale. This is why we would need a measurement that can be applied across different types of rewards. This measurement is the concept of "utility".

We should note the conceptual difference between reward and utility.

(i) The certainty of the reward and the variability of utility

The size of the reward is expressed in objective terms, e.g. additional income of RMB1000, a job promotion, a workplace commendation etc. The concept of the reward is an independent one, and it does not arise from a comparison. For instance, an income of RMB10,000 is simply that, RMB10,000 to the recipient, no matter who the recipient is and no matter what the recipient's specific circumstances are.

When we talk about utility, however, it is determined by the extent to which the individual's needs have been met. It is an abstract idea of benefit size, and this benefit is often related to the recipient's current circumstances. The same reward can have different utility levels for different people or for the same individual across different situations. For instance, an expected income of RMB10,000 would mean different things to an impoverished family and a billionaire.

(ii) The objective nature of the reward, and the comparative nature of utility

As the reward has a specific value, the size of the reward is objective and does not change depending on the recipient's circumstances. For example, an individual was hired at a wage level of RMB5000 a month and then received a raise to RMB20,000 a month. For him, his initial wage would always be RMB5000. He would not say that his starting wage "was not RMB5000" simply because he had received a raise since. On the other hand, utility is a concept that arises from comparison (in this volume we have adopted the concept of ordinal utility from the field of economics), which makes it difficult for us to state a specific "utility value". For instance, in a group of college graduates, some would find the utility of "going into research" greater while others would believe that "joining the workforce" presents greater utility. However, there is no way of accurately stating the "utility differential" between the two options.

(iii) The multivariate nature of the reward and the singular nature of utility

For a given behavior, the reward may come in a number of forms: economic reward, a reward in improved social standing etc. However, no matter the type of reward in question, there is only one kind of utility.

4.1.7.3 Exchange Utility as an Objective Measure for the Comparison of the True Value of Different Types of Rewards

Exchange effect, a way to estimate utility size based on the principle of exchange, is a method of comparing different rewards.

This method does not require the calculation of reward utility. Instead, it determines utility size based on whether individuals are willing to exchange one for the other.

If an individual is willing to exchange reward r_1 for reward r_2, then for this person the utility of reward r_1 is lesser than the utility of reward r_2. Therefore, $u(r_1) \leq u(r_2)$.

If the actor is willing to exchange reward r_1 for reward r_2, and is then willing to exchange reward r_2 for reward r_1, for this actor the utility of reward r_1 and reward r_2 are equal, or $u(r_1) = u(r_2)$.

For instance, a hungry person (Person A) is willing to spend RMB1000 to buy a loaf of bread from another person (Person B). For Person A, the utility of the loaf of bread is greater than the utility of the RMB1000. For Person B, he is willing to sell the loaf of bread for RMB1000, which means that to him the utility of RMB1000 is greater than the utility of a loaf of bread.

If an individual is willing to spend RMB1000 on a loaf of bread, and is also willing to exchange the loaf for RMB1000, then it means that for this person the two have equal utility.

Actually, on the commodities market, for the buyer the utility of the good purchased is greater than the effect of the coin paid for the purchase, while for the seller the utility of the coin received is greater than the utility of the good sold. The simultaneous existence of these two situations forms the basis for the occurrence of trading behavior.

4.2 Resource-Type Promoters

The existence of behavior resources is one of the necessary conditions for the occurrence of a given behavior. Resource-type promoters are a type of institution component which induce desirable behavior by providing the resources required for the occurrence of such behavior.

4.2.1 Resource-Type Promoters Provided Gratis

Resource-type promoters provided *gratis* are a kind of institution component designed to promote desirable behavior with the requisite resources provided no charge to the recipient.

Example 4.7 Research sponsorship by the National Natural Science Foundation of China

Scientific research can help spur social development, enhance the country's capabilities, and raise the general standard of living. Hence, the conduct of scientific research is desirable behavior. The establishment of the National Natural Science Foundation of China was approved by the State Council in February 1986. The National Natural Science Foundation mainly funds basic research and has played a key role in the promotion of basic research in the natural sciences in China.

Example 4.8 The sponsorship of entrepreneurship by the Shanghai Technology Entrepreneurship Foundation for Graduates

The Shanghai Technology Entrepreneurship Foundation is China's first instance of a not-for-profit organization seeking to promote entrepreneurship among college graduates. Once enterprises put in a grant application have been found to have met scheme requirements, the Foundation would then provide financial support up to RMB500,000 to the successful applicant. The college graduates who have applied for such funding would also be required to inject their personal funds into the enterprise at the same time, with the Foundation holding no more than 50 % shareholding in the funded enterprise. In addition, the Foundation would not receive any dividends or interest from the enterprise. Two years later, the college students successful in operating their businesses would be able to buy back their shareholdings at the original price. If the business fails, the investment would be written off by the fund and the entrepreneurs would not have to bear liability [for the funds invested].

4.2.2 Resource-Type Promoters Provided at a Cost

Resource-type promoters provided at a cost are a kind of institution component designed to promote desirable behavior with the requisite resources provided at a charge to the recipient.

Example 4.9 Special loans

Special loans are bank loans with a designated purpose. Generally speaking, special loans are used to support projects encouraged by the state such as for the economic development of rural areas and poverty eradication efforts. Special loans often come with preferential interest rates. The borrowing unit applies for the loan from the specialized bank, and following application approval by the specialized bank another application is then put in with the People's Bank of China. The loan is approved and given by the People's Bank of China in accordance with prevailing lending policies.

Example 4.10 The R&D Public Service Platform

The Shanghai R&D Public Service Platform is a resource and service institution designed to support the R&D efforts of enterprises. It offers users several amenities, such as the sharing of scientific data, a literature look-up service, the sharing of instruments and other equipment, collaborations for laboratory base etc. The aim of providing these services is to enhance enterprise innovation and to lower the cost of innovation. The Shanghai municipal government provides cost subsidies for enterprises using the instrument and equipment made available through the R&D Public Service Platform. A 30 % subsidy is provided for equipment fees up to RMB50,000, with an additional subsidy of a maximum of 15 % given for equipment fees exceeding the first RMB50,000 should total fees cross this level.

4.2.3 Combining Resource-Type Promoters and Observers

The aforementioned two types of resource-type promoters are often used in tandem with observers. For instance, the National Natural Science Foundation reviews research applications prior to the awarding of funding. Only projects that have been approved by the Foundation would receive Foundation funding. Applications for funding from other entrepreneurship funds and for special loans etc., must also be reviewed and approved before financial support can be provided.

4.3 Opportunity-Type Promoters

The existence of behavior opportunity is one of the necessary conditions for the occurrence of a given behavior. Opportunity-type promoters are a type of institution component which induce desirable behavior by providing the opportunities required for the occurrence of such behavior.

In a situation where behavioral utility and behavioral resources both already exist, behavioral opportunity becomes the determining factor for the occurrence of desirable behavior. It has often been said that hard work is useless without the requisite opportunity for one to showcase his talent. Therefore, the key function of the opportunity-type promoter is to provide opportunities for desirable behavior within the institution.

For example, to encourage the employment of college graduates, local governments have organized internship and job opportunities in various enterprises for fresh graduates on a regular basis. This practice is effectively an opportunity-type promoter for “graduate employment”.

To take another example, in a bid to nurture younger leaders, certain government organizations have created opportunities for them to “establish work performance”. Very often, these young people are seconded to the grassroots level where more

problems exist so that they would have the "opportunity" to show what they are capable of. The aim is to have the truly talented stand out from the rest of the pack in the course of actual work.

In addition, some enterprises also hold open speech competitions for employees seeking to become management members. This provides an opportunity for those with ideas but no room for showcasing their ideas to "show off their individual capabilities".

Some local governments have created "Entrepreneurship Information Platforms" targeted at college graduates. These are in effect a type of opportunity-type promoter as through such platforms fresh graduates would be able to look for market opportunities for their "entrepreneurship" behaviors.

References

Ma W (2014) State council allocates RMB10 billion to address atmospheric pollution, to substitute subsidies with rewards. China Times, 15 Feb 2014

Ye K (2001) Experiences in an options institution for entrepreneurs, theories and policy choices. Acad Bimestris 2001(1):57–64

Chapter 5
Behavior Suppressors

Behavior suppressors, or simply "suppressors" here, are institution components that are used to prevent the occurrence of target behaviors or to lower the effort level for such behaviors. Suppressors inhibit behaviors by removing the key conditions of the behaviors in question viable. Within the institution, the suppressor is used to induce individuals to abandon undesirable behavior or to reduce the effort level for such behavior.

There are two key elements in the suppressor: the first is the set of response rules associated with the suppressor, and the second is the implementing agency or individual who carries out the rules. In practice, for the implementation of suppressor response rules, the suppressor is often paired with an observer and typically integrated into a single mechanism to minimize costs and raise efficiency levels for the implementing agency. Therefore, in practice we often see "institution components" that comprise the integration of observer and suppressor, such as with "behavior measurement + suppressor response rules + observation and suppressor implementer". An example of such an institution component would be "observation of corrupt behavior by officials (e.g., expenses far greater than typical income) + penalties for corrupt behavior + implementing agency for the observation and penalizing of corrupt behavior (such as the Central Commission for Discipline Inspection or the courts)".

Behavior suppressors can be divided into the following types: reward-type suppressors, resource-type suppressors, opportunity-type suppressors, and cost-type suppressors.

As this chapter will deal with two types of suppressors, the reward-type suppressor and the cost-type suppressor, the difference between the reward and behavioral cost should be emphasized here. Behavioral cost typically arises at the same time as the behavior, as long as the latter occurs the former will also be incurred. Therefore, in general the issue of probability does not arise with the matter of behavioral cost, while behavioral reward generally occurs following the behavior. Following the occurrence of the behavior in question, very often we can attach a probability to the subsequent occurrence of the expected reward. This is why the concept of "expected reward" exists.

S. Sun, *Five Basic Institution Structures and Institutional Economics*,
DOI 10.1007/978-981-10-0344-8_5

Cost-type suppressors are a common type of institution component and come in two main sub-types: the first is a cost-type suppressor that is naturally formed. This sort of suppressor is determined by the characteristics of the behavior in question, and no matter whether the behavior is desirable or undesirable there is a cost associated with the occurrence of the said behavior. The second class is a cost-type suppressor that exists by the human design, and this sort of suppressor is used to curb undesirable behavior. For instance, as cigarette smoking is harmful to human health—which makes it an undesirable behavior—certain countries have chosen to impose hefty taxes on cigarettes in order to increase the cost of "smoking behavior" with the objective of curbing the act of smoking in mind.

5.1 Reward-Type Suppressors

Reward-type suppressors work by inflicting negative reward on undesirable behavior, with the link between negative rewards and the behavior in question made clear beforehand. This way, individuals targeted for management will expect the negative reward prior to committing such behavior and may be dissuaded from doing so as a result.

The accurate observation of behavioral state is critical in determining the effect of the reward-type suppressor. Therefore, when utilizing reward-type suppressors, specific attention must be paid to the selection of observers with strong performance to be used in tandem.

5.1.1 Suppressors for Economic Rewards

Suppressors aimed at reducing economic rewards link "economic reward" and "the measurement of undesirable behavior" together. How this type of suppressor works is: the graver the undesirable behavior observed, the greater the negative economic reward the actor receives.

Example 5.1 Poor performance of suppressor designed to curb pollution by enterprises

China's Rules for the Implementation of Water Pollution Control Act in tandem with the regulatory agencies tasked with the job of curbing water pollution serve as the suppressor for environment pollution by enterprises within Chinese borders. However, as the performance of this suppressor is poor, pollution by enterprises was not curbed effectively.

In general, there are two performance flaws with this type of suppressors:

(i) Negative reward too insignificant to constitute sufficient expected negative rewards for enterprises with regard to polluting behaviors.

In China's Rules for the Implementation of Water Pollution Control Act, it is stipulated that:

Enterprises that discharge toxic and polluting effluent [into the environment] are subject to a fine of up to RMB50,000, while enterprises found responsible for an incidence of water pollution are subject to a fine of 20 % of the direct damage caused with a limit of RMB200,000. Enterprises that have caused severe economic loss are subject to a fine of 30 % of the direct damage caused with a limit of RMB1 million.

For example, an enterprise that has caused a major pollution incident is fined the maximum amount of RMB1 million.

The half-year profit of this enterprise is RMB450 million, which means that it may earn RMB900 million for the entire year. For an enterprise that makes RMB900 a year, the fine of RMB1 million is but one-nine hundredths of its annual profit. Here, we see that even the maximum applicable penalty can be actually rather negligible. Hence, we can imagine the kind of "impact" that lesser fines of RMB50,000 or RMB100,000 would make.

(ii) Enforcement agencies do not possess sufficient power in terms of observation and implementation.

In China, agencies legally empowered to deal with the problem of pollution appear to be weak. For instance, the US Environmental Protection Agency is a mega-agency that employs over 17,000 employees and has an annual operating budget of more than USD10 billion a year. In contrast, China's State Environmental Protection Administration employs just over 200 public servants. We can hence imagine how effective these agencies are in terms of penalizing errant enterprises.

Another source of the corrosion of these agencies' regulatory effectiveness lies in the inaction of agency officials. Reports have it that, in a certain incident when an enterprise had discharged polluting effluent into a river, letters were written to the head of the province, who then forwarded the correspondence to the municipal environmental protection bureau. Bureau personnel were then sent to the enterprise in question twice to ascertain the situation. However, these personnel had kept mum despite witnessing the situation at first hand.

5.1.2 Suppressors for Reputational Rewards

This class of suppressors links the "lowering of one's reputation" with "undesirable behavior measured". How this type of suppressor works is: the graver the undesirable behavior, the greater the negative reputational reward the actor receives.

For example, bureaux of quality and technical supervision have been set up across China to conduct regular sampling of products found in local markets for quality supervision purposes. These bureaux then publicly disseminate lists of enterprises found to have manufactured sub-standard products (this is a negative reputational reward). These institutions effectively act as suppressors for the "manufacturing of sub-standard products" by enterprises. Enterprises that appear on

these lists see demand for their products drop as a result with significant economic loss caused (this is a negative economic reward). If an enterprise recognizes the connection between these outcomes, it would work to ensure the quality of its products to avoid the production of sub-par items.

Example 5.2 The annual 315 or 15 March Gala on China Central Television (CCTV) is, in effect, a reputational-type suppressor for the "piracy and manufacturing of imitation goods" by enterprises

The annual 315 Gala is a large-scale public-service event held jointly by CCTV and organs and agencies such as the Supreme People's Court, the Supreme People's Procuratorate, the NPC Standing Committee, the Legislative Affairs Commission, the Ministry of Industry and Information Technology, the Ministry of Public Security, the Ministry of Justice, the Ministry of Agriculture, the Ministry of Commerce, the General Administration of Quality Supervision, Inspection and Quarantine, the Legislative Affairs Office of the State Council, the State Intellectual Property Office, the Food Safety Commission Office of the State Council, the State Food and Drug Administration and the China Consumers' Association. The aim of the event is to protect the rights and interests of consumers. The 315 Gala has been broadcast live across China since 1991.

For over 20 years, the 15 March event has lifted the veil on numerous scams and other shady practices, leading to resistance of pirated and poor-quality products on the part of the Chinese people. As such, it can be said that the program has dealt severe blows to various illegal business operations.

Commentary:

Through tips and reporting by its journalists, CCTV has acted as an observer of the "piracy and manufacturing of imitation goods" by enterprises. With CCTV's broad reach, the program also enhances negative reputational rewards to enterprises featured as consumers across the country are made aware of, and subsequently avoid, such products. This in turn generates massive negative economic rewards for such enterprises, which makes this type of suppressors rather effective.

5.1.3 Suppressors for Rewards Pertaining to Social Standing

This class of suppressors links the "lowering of social standing" with "undesirable behavior measured". How this type of suppressor works is: the graver the undesirable behavior, the greater the actor's social standing falls. For instance, the penalty for officials engaging in corruption would be "punitive demotion" in various degrees. The stipulated penalty here and corresponding implementation mechanisms together constitute a reputation-type suppressor with regard to corrupt behaviors.

5.1.4 The Integrated Use of Various Types of Reward-Type Suppressors

In practice, for the best deterrent effect, a variety of reward-type suppressors are deployed in tandem. For instance, laws in various countries allow for the concurrent fining and imprisonment of criminal offenders who have been found guilty. This sort of penalty comprises several aspects: economic loss, opportunity loss, reputational loss and bodily suffering. At this juncture, the law (or rather, the association between negative rewards and the undesirable behavior in question) and the enforcement agency form an integrated suppressor against illegal behaviors.

5.2 Opportunity-Type Suppressors

Opportunity-type suppressors are a type of institution component and work by removing (in part or in whole) the behavioral opportunity for the undesirable behavior in question to prevent the occurrence of the said behavior or to reduce the probability that it would happen.

5.2.1 Opportunity-Type Suppressors that Work Through an Admittance Mechanism

Admittance institutions and the relevant administrative agency may constitute an opportunity-type suppressor. An admittance institution is an administrative institution that requires individuals (enterprises or persons) to show that they meet qualifying standards before they are allowed to engage in specific behaviors. Through this method, individuals without the requisite admission qualifications will lose the opportunity to select the relevant behavior.

Example 5.3 The food production permit institution for food manufacturing enterprises

The food production permit institution for enterprises together with the relevant management mechanisms can constitute an opportunity-type suppressor for “the production of non-compliant food products” in practice.

To prevent food manufacturing enterprises from producing non-compliant food products that threaten public health, the Chinese government has implemented a food production permit institution as a way to manage the issue. The corresponding supervisory agency reviews the production environment, equipment, raw materials used, processes, product standards, technician qualifications, storage safety level, detection management level, packaging situation etc. at the food manufacturer’s and conduct test sampling of the manufacturer’s products to evaluate the enterprise’s

overall food safety standard. Food manufacturing enterprises evaluated and deemed to have met requirements are then issued with food safety and manufacturing permits. Legally speaking, only enterprises with such permits are allowed to manufacture food products.

The state has established the food production permit institution as well as the corresponding supervisory organs in order to curb the manufacturing of food products by enterprises that do not meet its requirements. However, currently the effect of this type of suppressor is not very high as those empowered to implement the institution do not have enough power in an environment rife with smaller enterprises with production facilities that are plainly not up to scratch. This means that regulatory enforcement teams are overwhelmed in such situations.

Example 5.4 The Certified Public Accountant (CPA) institution

The Certified Public Accountant institution along with its accompanying management institutions constitutes an opportunity-type suppressor for "the performance of auditing work by less-qualified individuals".

Under the law, accounting firm personnel who perform audits are required to hold a CPA qualification. In addition, to obtain a CPA qualification, the candidate has to sit for a rigorous examination of his/her knowledge of accounting as well as undergo a background check. Individuals who have received penalties for criminal offenses, individuals who have received administrative penalties for severe errors in finance, accounting, auditing, corporate management or economic management, and individuals who have been dismissed from their job positions or have got harsher penalties are not eligible for CPA registration.

Example 5.5 The home purchasing restriction policy

In 2011, to prevent a housing bubble, the Chinese government introduced a policy which, together with the relevant implementation policies from the agencies concerned, restricted individuals from the undesirable behavior that was "the purchase of multiple residential units". For example, the Shanghai municipal government mandated that families with *hukou* status in the city and owning two or more residential units were no longer allowed to purchase more residential units in the city. The same restriction was also applied to: households not holding Shanghai *hukou* registration and households unable to supply evidence of tax payment in the city for one year or more and which already own one or more residential units in the city. Following the introduction of this policy, a number of people lost the opportunity to purchase residential units in Shanghai.

5.2.2 Opportunity-Type Suppressors that Work by Masking or Hiding Information

Information masking institutions and the relevant implementation agency constitute an opportunity-type suppressor that work through the hiding of information.

Information masking works by concealing certain information from actors who may commit undesirable behavior in order to curb the said undesirable behavior.

Example 5.6 The use of candidate numbers in place of names on examination papers

In China's *gaokao* or college entrance exams, to prevent undesirable behavior such as evaluator bias, exam candidates are required to enter their candidate numbers in place of their names on exam papers. This method is a type of opportunity-type suppressor as it eliminates opportunities for cheating by evaluators.

Example 5.7 Information masking as a means against smuggling

Smugglers typically facilitate their work by bribing customs officials in advance. The situation usually goes like this: the smuggler bribes customs officials he is familiar with, and said officials would advise the smuggler of when and where he could expect the official to be on duty. At the appointed time and location, the smuggler then tries to move his goods across the border. This way, the smuggler could be more confident about not getting caught in the act. To curb the undesirable activity that is "the bribery of customs officials", customs authorities at a certain border crossing have implemented a new institution with over 200 customs officials. Within the new institution, officials would only know their specific post locations after clocking in and 5 min before they are to assume their posts. In addition, these officials are barred from bringing mobile phones to work. This way, smugglers would not be able to know where their "preferred" customs officials would be in advance and hence lose access to their "safe channels" of passage. Following the implementation of the institution, the number of smuggling cases at the said border crossing was drastically reduced.

5.3 Resource-Type Suppressors

Resource-type suppressors are a kind of institution components used to inhibit the undesirable behavior by depriving them of resources.

Example 5.8 The prohibition on firearms

The ban on firearms and corresponding enforcement agencies together form a resource-type suppressor with regard to crime.

Firearms are a crucial resource for the committing of crimes, and in order to curb criminal behaviors the Chinese government has established stringent control of firearms.

The Law of the People's Republic of China on Control of Guns stipulates that the government shall maintain stringent control of firearms. The possession, manufacturing, transport, leasing and lending of firearms by any unit or individual are strictly forbidden under the law. China deals strictly with any breach of its laws on the firearm restriction.

The public security agencies are responsible for managing the restriction of firearms in China.

The Criminal Law of the People's Republic of China stipulates that individuals found guilty of contravening firearms laws with illegal possession of firearms or explosives may be imprisoned in jail for up to three years, detained or otherwise have his or her movements restricted; in severe cases, the offender may be jailed for a period between three and seven years.

Example 5.9 Resource-type suppressor designed to curb "drinking by college students"

A certain college has discovered that some of its students have been drinking on campus, with undesirable consequences such as the frittering away of student allowances on alcohol, the occurrence of incidents, and negative impact on students' academic performance. Hence, the college has expressly forbidden the drinking of alcohol by students in its rules. This is a reward-type suppressor aimed at curbing student drinking. Here, the regulations and penalties surrounding student drinking are the response rules, while the implementing personnel are the counselors. In practice, however, this suppressor has not worked well. First of all, observation is a challenge, as there are any number of locations across campus where students may indulge in drink and there is no way counselors can keep an eye on all students round the clock. Therefore, it was difficult to uncover cases of student drinking. What this means is that the observer paired with the suppressor is performing poorly. Second, penalties associated with the behavior are comparatively light as the drinking of alcohol is not a severe mistake. Generally speaking, the penalty for student drinking is a session of "conversation and criticism", which is essentially a rap on the knuckles for students involved. This means that the rules associated with the suppressor lack strength. Hence, this suppressor did not work very well, and student drinking continued in mockery of the rules. Some students would purchase alcohol as soon as they received their bursary allowances, and would then try to borrow money from their schoolmates or even lecturers when the money ran out.

Later on, the school found a solution to the problem: bursary allowances to students were loaded directly onto meal cards, the balance of which could not be converted to cash. These meal cards were only accepted in the school's cafeterias where alcohol was not sold. This way, some students were unable to purchase alcohol as their bursary grants could only be used for their meals. The problem of student drinking was curbed significantly as a result.

Here, the "loading of student bursaries directly onto meal cards" and "no sales of alcohol in campus cafeterias" etc. act as suppressive rules while the institutions for bursary management, meal card management, cafeteria management etc. act as institutions that implement such rules. Together, a resource-type suppressor that acts to curb "student drinking" is formed. For the behavior that is "student drinking", this resource-type suppressor is more effective than a reward-type suppressor as measures such as "loading of student bursaries directly onto meal cards" and "no

sales of alcohol in campus cafeterias" can be easily implemented and performed well.

From this example, we can see that the impact of different suppressors can vary widely. Therefore, the selection of the appropriate suppressor is important during institution design.

5.4 Cost-Type Suppressors

5.4.1 Cost-Type Suppressors that Have Formed Naturally

The majority of cost-type suppressors are formed by non-institutional factors. This type of suppressors is known as "cost-type suppressors that have formed naturally" or simply "natural cost-type suppressors". The response rules associated with this type of suppressors tend to follow certain patterns (for instance, product cost increases in a linear fashion as production scale increases) and may hence be described using functions.

For example, with the cost function of "fixed + variable costs" in production which we commonly see in economics, the variable component (commonly in the form of effort level and typically manifesting as the enterprise's production scale) can be expressed in a positive linear or non-linear function. When production scale is extremely great, fixed cost can be ignored and the enterprise's production cost is only related to the scale production. At this point, the linear cost function looks like this:

$$c_i = e_i$$

In the equation, the enterprise i has a cost of c_i, $c_i \geq 0$, with production scale at e_i, $e_i \geq 0$.

Together with the "Output Function-Type Reward Promoter" (see Chap. 4) designed by the author, the cost function plays an important role in the theoretical analyses of institution structures in Chaps. 9 and 10.

5.4.2 Designed Cost-Type Suppressors

The designed cost-type suppressor comes from designing artificially and is an institution component used to inhibit the undesirable behavior by increasing its cost.

For instance, border troops and various defensive facilities put in place along borders by various nations constitute a cost-type suppressor for any enemy wanting to start an "invasion of another country". If an aggressor were to act in such a way, a high price would have to be paid. Certain enterprises have adopted anti-burglary

measures such as the installation of security doors and the erection of boundary fences and walls. All these are cost-type suppressors for "burglary", as would-be burglars would have to put in a higher cost during the act due to the existence of these measures.

5.5 Suppressors by Drive Type

The reward-type suppressor, resource-type suppressor, opportunity-type suppressor and cost-type suppressor analyzed above are categorized by the behavioral conditions targeted by the suppressor.

If we are to categorize suppressors by drive type, they can be divided into automatic-response suppressors and response suppressors by design.

5.5.1 Designing an Automatic-Response Suppressor

Automatic-response suppressors are suppressors that realize response rules automatically for reasons of nature or social norms. For example, a person who does not pay attention to matters of hygiene while cooking at home will cause food poisoning to himself and his family. This is essentially an automatic reward-type suppressor.

In the course of institution design, it is best to try to use such suppressors as much as possible to inhibit undesirable behavior. This is because automatic suppressors are "natural" suppressors and typically do not incur management costs.

Example 5.10 Suppressing "the unequal distribution of gruel"

In Sect. 1.1 of Chap. 1 in this volume, we looked at the "institution of gruel distribution" where a group of seven individuals shared a bucket of gruel among themselves each day. However, no matter who acted as the "gruel distributor" for the day, "the unequal distribution of gruel" would always happen, with the person distributing the gruel reserving a larger portion for himself. Finally, the solution settled upon was: "the one who distributes the gruel gets his share last". In other words, the person responsible for the distribution of the gruel may only get his bowl after everyone else has made their selections. This way, the person in charge of gruel distribution would work to avoid getting the bowl with the least amount of gruel by making sure that the gruel is equally divided across every bowl. Thereafter, there was no more conflict within the group, and everyone lived happily ever after.

Let us now think for a moment: what had suppressed "the unequal distribution of gruel"?

We can see that the rule that "the one who distributes the gruel gets his share last" has made for an automatic-response suppressor. If the individual responsible for dividing up the gruel for the day does so unevenly, then everyone would fight

for "the bowl with a little bit more gruel in it" and the last person (i.e., the distributor) would be left with the bowl "with the smallest share of gruel". This way, there is a "negative reward" attached to the undesirable behavior that is "the unequal distribution of gruel". At this juncture, if he wishes to eat a little more, he would have no choice but to ensure the equal distribution of gruel across every bowl.

Reward-type suppressors are but one type of automatic suppressors. Take the example of a parent reducing his or her child's pocket money upon the realization that the child has been skipping classes to visit the Internet cafe, so that the child would no longer be able to pay for hours at the cafe. From the perspective of the school seeking to manage students' behavior, this is an automatic resource-type suppressor. If parents are to come together to petition the government for the closure of all Internet cafes in the vicinity of the school, such a closure would deprive their children of the opportunity to visit Internet cafes. From the perspective of the school seeking to manage student behaviors, this actually constitutes an automatic opportunity-type suppressor for "using the Internet at the Internet cafe".

In practice, in general for automatic suppressors the implementation and enforcement or response rules may be performed by those who are negatively impacted by the said undesirable behavior. They are most incentivized to implement the "suppression rules" as their interests are threatened by the undesirable behavior in question. For instance, if an environmental protection agency is shorthanded, it may choose to hire a local resident of the area affected by pollution as the latter would be even more motivated to enforce environment protection laws.

5.5.2 Designing a Response-Type Suppressor

A suppressor response is an institution component that is designed in line with response rules set by the designer to have a suppression effect on a certain behavior. Examples 5.1 to 5.10 in this chapter all feature the use of response-type suppressors. This type of suppressor exists by human design and typically requires an implementing agency or individual to ensure the rules are being enforced. Therefore, the management costs associated with such suppressors are comparatively higher.

5.6 Suppressor Failures

The term "suppressor failure" refers to the situation where the suppressor is no longer effective at curbing the targeted undesired behavior. In terms of the reasons for suppressor failure, there are two main types: limit failure and compensatory failure.

5.6.1 Limit Failures

Limit failures are instances where the conditions targeted by the suppressor (such as rewards, resources, and opportunities) are at their limit and cannot be taken away further, leading to the loss of suppression performance. In such situations, only the "haves" would be affected by the suppressor and not the "have-nots".

For example, in the instance of the reputation-type suppressor mentioned earlier, it would only work with larger enterprises with a reputation to protect. As for smaller operations with poor reputations to begin with, the threat to one's reputation would simply not matter as much. In China, many *shanzhai* or pirated/imitation goods compete on the basis of "extremely low prices". The vast majority of such products are known to the consumer as *shanzhai* prior to purchase, which means that the seller's or manufacturer's reputation is not at stake here.

As limit failures can occur with behavior suppressors, when designing and deploying suppressors it is important to examine if the targets for behavior management possess the requisite conditions that may then be taken away. Second, penalties should be cautiously meted out in appropriate degrees; if the maximum penalty were to be meted out too casually the deterrent effect would be lost for the future.

Example 5.11 The negligible impact of "voiding of exam results" and "three-year ban from exams" for poorer performers found to have cheated on graduate school admissions exams

To tackle the problem of candidates cheating during graduate school admissions exams, the Ministry of Education of China has mandated that a student caught cheating would have his or her exam results voided. In addition, the student would be banned from sitting for the exams for the next three years.

At first glance, this rule may seem extremely harsh. However, as a punitive threat it has very little impact on students who perform more poorly.

A key question here is: what is the profile of a "cheating" student like?

Generally speaking, a better performer should have no problems making it to the graduate school even if he or she does not resort to cheating at the exams. In addition, with cheating there was also the risk of having one's results voided if one were to be caught in the act.

On the other hand, students who perform more poorly academically would have no hope of making it to graduate without resorting to cheating. For these students who otherwise would not be able to gain graduate school admission via the normal route, there is no further loss to them if their exam results were to be voided or they were to be banned for the exams for three years. On the other hand, if the cheating is not discovered, the candidate will have a chance of making it to a graduate program. For such students, since their chances of making it to the graduate school would be hopeless without recourse to cheating, clearly, their "optimal behavior" here would be "cheating".

5.6.2 Compensatory Failures

Compensatory suppressor failures occur when additional compensation that cannot be controlled exist outside of the element that the suppressor is designed to remove. At this point, the performance of the suppressor would be significantly affected.

For instance, for an economic-type suppressor, if individuals targeted for management have other sources of income not affected by the suppressor, then the suppressor is unlikely to work well. If income other than what is affected by the suppressor is sufficiently large, then individuals targeted for management may not even be bothered by punitive measures such as the "withholding of bonuses". At this juncture, the use of economic-type suppressors would no longer be effective in curbing undesirable behavior.

In another example, the state's "prohibition on firearms" policy (a resource-type suppressor) has made it impossible for individuals to purchase firearms in regular stores. This way, the suppressor has had quite a good effect in terms of curbing violent crime. However, a black market for firearm exists, which would affect the performance of the suppressor to some degree.

Due to the phenomenon of compensatory failures, when selecting suppressors during institution design, attention must be paid to whether there is any compensation mechanism involved with elements targeted for removal using the selected suppressor(s).

Chapter 6
The Sun Diagram and the Five Basic Institution Structures

6.1 The Significance of the Sun Diagram

The Sun Diagram is a diagram made up of symbols. It was developed by the author as a way to describe institution structures, and its full name is the "Graphic Symbol and Rules System for Institution Design". As this name is rather long and unwieldy, it has also been christened the "Sun Diagram".

The significance of the Sun Diagram lies in the following few areas.

First, the Sun Diagram illustrates the institution structure clearly and concisely.

We can say that issues in management institutions have always been an important topic both in management science and in economics. However, over the years, there have been very few successful studies of management institutions which have rendered good results. The key reason for this situation is the lack of scientific analytic tools for institution analysis and common design tools and techniques, with the vast majority of studies relying on personal experience and thinking. This means that the resulting studies lack depth and are flawed. Without the use of an overall institution or institution diagram annotated with clear symbols, an over-reliance on intuition and experience can make it difficult to observe the multivariate factors in, and complex structure of, an institution or institution. Under such circumstances, scientific calculations cannot be made, making it difficult for a precise comparison and consideration of effects and costs of various institution structures. A variety of solutions have been proposed for similar problems plaguing the institution, and this is precisely because these solutions have been forged in individual, personal experiences.

The Sun Diagram was designed to address the above issues. It can be regarded as the "circuit diagram of institution design", with various symbols used to represent institution "components" and line segments used to connect various institution elements in what is a clear and complete illustration of institution structure and the relationships between various institution elements.

S. Sun, *Five Basic Institution Structures and Institutional Economics*,
DOI 10.1007/978-981-10-0344-8_6

Second, the Sun Diagram allows institution design to be carried out in a simple, straightforward manner like with the engineering design.

As the Sun Diagram is an accurate and concise way of expressing the institution structure, during the institution design process the designer would no longer have to imagine how the institution would work, just like a blind man would have to imagine how a certain gadget described looks like. Instead, the designer would be able to use the Sun Diagram to describe the institution just like an engineer [can use a drawing to describe a piece of machinery] and use the Diagram to analyze and resolve existing problems within the institution (i.e., institution design).

Third, with the Sun Diagram we can identify and classify various types of management structures.

In management practice, there are numerous management institutions. Generally speaking, there is no concise tool in use that can accurately present the various characteristics of an institution. As such, it is difficult to sort and classify the kinds of institutions that are in existence. The Sun Diagram is an abstract way of describing the management institution. Using the Diagram, we can lay out the basic features of various institutions and discard non-critical or peripheral institution characteristics at the same time. This way, we will be able to see that certain institutions from various areas of applications fundamentally share similar structures as well as see how various institutions differ from each other in fundamental ways. As such, we will be able to sort and classify various institution types as we gain a closer and better understanding of these institution types.

Later on, we will see that in reality, the institution structures that are in existence can all be classified as one five basic types following analysis with the Sun Diagram. This conclusion is a key discovery of this volume.

6.2 Definitions of Various Terms Used in the Sun Diagram

The Sun Diagram comprises the behavior set, line segments and institution components.

6.2.1 The Actor

The term “actor” refers to individuals targeted for behavior management within an institution and who share common characteristics and interests. An example of institution actors would be the candidates within a given exam institution. If the designer feels that these individuals are of a similar capability level, then they can be regarded as one group of individuals targeted for management, or a single actor. However, if the candidates are categorized as “outstanding students”, “average

performers" and "poor performers", then together they constitute a total of three actors within the institution.

We can sort and classify institutions described using the Sun Diagram by the number of actors within the institution. Institutions with only one actor are known as single-actor institutions while institutions with more than one actor are known as multiple-actor institutions.

6.2.2 Behaviors and Behavior Set

Within the management institution, the actor is able to make behavioral choices based on his judgment of behavioral utility. The range of possible behaviors that a certain actor may opt for constitutes the actor's behavior set in this situation. For instance, exam candidates can choose to take the exam "normally" or to cheat during the exams. Here, "taking the exam normally" and "cheating at the exams" are both behaviors, and these two behaviors together make up the actor's "behavior set".

Behavior sets can be categorized as "discrete" or "continuous".

In a discrete behavior set, the elements are the various relevant behaviors b_i. There are at least two possible behaviors for the actor to choose from in a discrete behavior set. The behavior set must be complete and contain mutually-exclusive elements. This means that the actor must choose one of the behaviors within the behavior set (completeness) and can only opt for one behavior (mutually-exclusive). In a single-actor institution, the subscript i indicates the various behaviors that can be chosen by the actor. In a multiple-actor institution, the subscript i generally comprises of two elements: a reference to the actor in question and a reference to specific behavior choices. For specific rules on subscript i labeling, please refer to "Numbering rules for behavior" in Sect. 6.3.

In a continuous behavior set, the element involved is the level of effort e_i put into a certain behavior. Values of behavior effort level are located along a continuous line. Take note that $e_i > 0$. Here, the subscript i typically indicates the specific actor in question within the institution.

In certain institution models, behavior b_i within a discrete behavior set also comes with an effort level e_i. Under such circumstances, the actor may either choose one of several behaviors b_i or opt for a specific effort level e_i for various behaviors. What the actor eventually opts for is determined by the specific institution in play. In such cases, there would be a note on the situation in the Sun Diagram.

6.2.3 Institution Components

The institution component is an apparatus used by the institution designer to control the behavior of actors under management. Except for the "consequence" as a type

of transient component, other components are called as promoters, suppressors and observers (which can be concluded as "XX 器" in Chinese).

6.2.4 Line Segments

Line segments are used in the Sun Diagram to connect various elements in the diagram. Line segments also indicate the relationship between various institution elements.

For the sake of simplicity, in general line segments comprise either of straight or broken lines sans arrow heads. However, in situations where there may be ambiguity, arrow heads can be used to indicate the direction of influence. Here, the element that the arrow is pointing to is the element that is being impacted, while the element where the arrow originates is the element that is exerting an influence.

6.3 Symbols for Institution Components and Behavior Set in the Sun Diagram

6.3.1 Consequences (Outcomes)

"Consequences" are abbreviated as "con" by the author in the Sun Diagram.

In the Sun Diagram, very often we have to indicate certain consequences following the choice of a specific behavior by the actor. For example, when multiple individuals act in concert (cooperate), the cooperation would result in an outcome (e.g., in the case of trying to lift a heavy object together). This outcome would then lead to a certain reward (such as wages). The symbol for "consequence" is as shown in Fig. 6.1.

Like other institution components, the consequence also often comprise response rules to the behavior (either an inherent rule or an institution rule) and implementors who enforce such rules (natural implementors, or by institution design). However, when such consequences have a direct impact on the behavior of the actor, it is in fact a behavior promoter or suppressor. At this point, we can indicate the element by using the concepts "promoter" or "suppressor". Only when the consequence does not have a direct impact on behavior is it expressed using the concept "consequence". Therefore, in the Sun Diagram, the "consequence" is a type of transient component.

Fig. 6.1 Consequences

6.3.2 Promoters

Promoters are used in an institution to encourage a behavior [into opting for a certain behavior]. Promoters are often deployed when the probability that a certain behavior would occur or the level of effort that goes into said behavior needs to be enhanced.

Within the Sun Diagram, there are three types of promoters: the reward-type promoter, the resource-type promoter, and the opportunity-type promoter.

6.3.2.1 Reward-Type Promoters

The reward-type promoter is used to enhance the probability that a certain behavior would occur or the level of effort that goes into the said behavior through the use of positive rewards. Examples of reward-type promoters include wages and job promotions.

The symbol for the reward-type promoter is as shown in Fig. 6.2, with the component represented by the letter "r".

6.3.2.2 Resource-Type Promoters

The resource-type promoter is used to enable the occurrence of a certain behavior or greater behavioral effort through the provision of behavioral resources. For example, to promote the behavior that is "research in the sciences", the Chinese government has often established "R&D grants" as a form of financial support for research activity. Here, the measure that is the R&D grant is in fact a resource-type promoter for "research in the sciences".

The symbol for the resource-type promoter is shown in Fig. 6.3, with the text abbreviation "res".

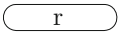

Fig. 6.2 Reward-type promoters

Fig. 6.3 Resource-type promoters

6.3.2.3 Opportunity-Type Promoters

The opportunity-type promoter enables the occurrence of a given behavior through the provision of behavioral opportunities. For example, the "*gaokao* college entrance exam institution" is effectively an opportunity-type promoter for "participating in college entrance exams" as only with the former will the latter have a chance of happening.

The symbol for the opportunity-type promoter is shown in Fig. 6.4, with the text abbreviation "opp" used.

6.3.3 Suppressors

Suppressors are used in an institution to diminish the probability that a given behavior would occur, or to decrease the level of effort that goes into the target behavior.

In the Sun Diagram, there are four types of suppressors that are commonly used: the reward-type suppressor, the resource-type suppressor, the opportunity-type suppressor, and the cost suppressor.

6.3.3.1 Reward-Type Suppressors

The reward-type suppressor works by bringing negative rewards to the target behavior, hence causing the actor losses should he opt for such a behavior. This in turn diminishes the probability of the target behavior occurring or the level of effort put into such behavior. Examples of reward-type suppressors include fines and job demotions.

The symbol for the reward-type suppressor is as shown in Fig. 6.5, with the component represented by the letter "s".

Fig. 6.4 Opportunity-type promoters

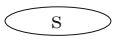

Fig. 6.5 Reward-type suppressors

6.3.3.2 Resource-Type Suppressors

The resource-type suppressor works by withholding behavioral resources to diminish the chances that the target behavior would occur or to diminish the level of effort put into the target behavior. An example of the resource-type suppressor is "the ban of the sale of firearms" as a means to "preventing armed robbery". This measure effectively removes the resource conditions that make armed robbery possible.

The symbol for the resource-type suppressor is as shown in Fig. 6.6, with component represented by the letters "res".

6.3.3.3 Opportunity-Type Suppressors

The opportunity-type suppressor works by eliminating the opportunities that make the target behavior possible or which enhance the level of effort put into the target behavior. For example, during a job performance review, the names of review committee members are kept confidential. This measure is an opportunity-type suppressor designed to diminish the chances of the behavior that is "bribery of committee members", as the former would diminish or even completely eliminate the opportunities that make the latter possible.

The symbol for the opportunity-type suppressor is as shown in Fig. 6.7, with component represented by the letters "opp".

6.3.3.4 Behavioral Cost Suppressors

Behavioral cost is inherently a behavioral suppressor. Thus, in the Sun Diagram it is expressed the same way as other suppressors. As we can see in Fig. 6.8, it is expressed

Fig. 6.6 Resource-type suppressors

Fig. 6.7 Opportunity-type suppressors

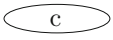

Fig. 6.8 Behavioral cost

as "c_i" in text. Note that the subscript "i" typically indicates the actor. When there is only one actor within the institution, the subscript notation is not required.

6.3.4 Observer or Classifier or Probability Device

In the Sun Diagram, the observer, classifier and probability device are fundamentally similar in nature. Therefore, they are all expressed as "$p_{i1}/\cdots/p_{in}$" in text (in the actual Sun Diagram, for the sake of simplicity, where there is no ambiguity they can also be expressed using "p_i", for instance the observer p_{21}/p_{22} can be expressed simply using "p_2"), with i generally the subscript for behavior b_i under observation. The distinct feature of the observer is that it has one input terminal and n output terminals (an observer with n output terminals is known as the "n-dimensional observer), with each output associated with a certain probability. These probabilities $p_i + p_{i+1} + \cdots + p_{i+n} \leq 1$ may not add up to 1 as outcomes with no significance may be omitted from the Sun Diagram. Hence, the output of the observer, classifier and probability device may not be complete (i.e., contain all possible outcomes). In the actual Sun Diagram, whether the outputs of these components are complete is determined by specific needs.

However, the outputs of some components are mutually-exclusive, and these components include observer, classifier and probability device. That is to say, there should only be one outcome for any given input. Take for instance the observer: with regard to the target behavior, observer outcome can only be either "behavior observed" or "behavior not observed".

The key difference between the observer, the classifier, and the probability device lies in their respective purposes or significance.

The observer is an apparatus deployed for the detection of behavior by actors under management, such as in the case of the invigilator during an exam. The observer's function is to detect behavior elements within a given discrete behavior set which have already occurred, with two possible outcomes: "behavior detected" and "behavior not detected". Each outcome occurs with a given probability (i.e., observer sensitivity or observation intensity, which is related to observer performance), and generally speaking [observation or detection] does not occur 100 % of the time (i.e., in all instances).

Hence, an observer for a discrete behavior set typically has one input terminal and two output terminals (i.e., a binary or two-dimensional observer). Figure 6.9 shows the graphic symbols for an observer used for a discrete behavior set, which is notated using the symbols "p_1/p_2" or p. Here, p_1 is the probability of "behavior detected" while p_2 is the probability of "behavior not detected".

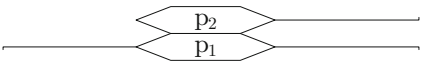

Fig. 6.9 Binary behavior observers

In reality, there are observers with more than two terminals or sides. This way, we can sort and classify observers based on the number of output terminals found on the observer. Observers with two output terminals are known as binary or two-dimensional observers, while observers with three output terminals are known as ternary observers, and so on and so forth. For example, when monitoring a specific behavior, if the possible outcomes of such monitoring are "behavior did not occur; behavior occurred albeit with low effort levels; behavior occurred with high effort levels", then this observer is a ternary observer.

The classifier is an apparatus that sorts or classifies a certain target. For example, the "examination" is what sorts students into the following categories: excellent, good, average and poor. Therefore, the "exam" can be seen as a kind of classifier. Indeed, the "exam" can also be understood as an observer, in that the aforementioned student categories are the result of an observation of their respective performances. As such, the exam can be regarded as a "four-dimensional observer". It is quite clear that the observer and the classifier are similar in nature. Indeed, the process of classification can only be performed through observation. Figure 6.10 is the symbol for a ternary classifier, "$p_1/p_2/p_3$".

At times, when the process is at a certain juncture, a certain outcome will emerge at a given probability level (and "will not emerge" will also occur at a complementary probability level). However, "will not emerge" would not have any significance within the Sun Diagram. Here, "this outcome does not occur" can be omitted from the output terminal, thus creating a "classifier" with only one output terminal. Take for example the exam scenario: the outcome of the exam can either be "pass" or "fail". However, if "fails exam" does not have a significance within the Sun Diagram, then this consequence can be omitted. As such, there will only be one output terminal, "pass" for the exam, and this output occurs at a given probability. This type of classifier (with only one output terminal) cannot be termed as a one-dimensioned classifier, because although it is illustrated with only one terminal in the Sun Diagram, it actually operates by using two dimensions or outcomes. In fact, institution components with only one output terminal cannot be termed as "classifiers" as the very definition of classification implies that there must be at least two outcomes. Instead, the component would be known as a "probability device". In Fig. 6.11 we see the symbol for the probability device.

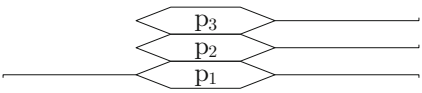

Fig. 6.10 Ternary classifier

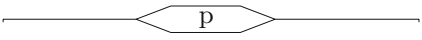

Fig. 6.11 Probability device

In the Sun Diagram, this type of probability device is often used to link the behavior and expected outcome. What it indicates is that the outcome would occur at a given probability (and not as a matter of fact) following the occurrence of said behavior.

6.3.5 Symbols for the Actor and Behavior Set

In the Sun Diagram, behavior elements in a discrete behavior set are typically notated using the symbol b_i, which indicates that it is the i behavior in the behavior set. In the case of a continuous behavior set, the behavior element is notated using the symbol e_i, which indicates that it is the continuous behavior set of the i actor (the element in a continuous behavior set is typically a continuous line of values for effort level, etc.). The meaning of the subscript differs as during the institution analysis process in Chaps. 7, 8, 9 and 10 later on we will see that most institutions with discrete behavior sets are single-actor institutions while institutions with continuous behavior sets are typically multiple-actor institutions.

In the institution analyses that follow, an actor is typically associated with one behavior set (although in reality an actor can be associated with more than one behavior set, as the institution being analyzed is usually concerned with only a certain aspect of the actor's behavior—such as production behavior or learning behavior—we have chosen to make this assumption). In the case of a discrete behavior set, the various behavior elements in a behavior set are laid out horizontally in a series of connected boxes, with each box standing for one element. Each box can be linked to other institution components from the top or bottom with the use of line segments.

Below, we look at the various symbols and text abbreviations for various behavior types.

6.3.5.1 Discrete and Continuous Behavior

i. Discrete behavior

A given number of behavior elements make up the behavior set for a certain actor, with symbols used shown in Fig. 6.12.

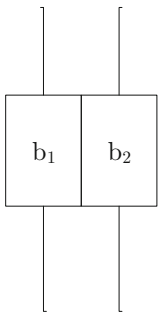

Fig. 6.12 Binary behavior for a certain actor

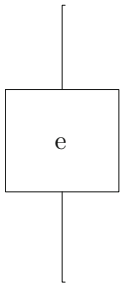

Fig. 6.13 Continuous behavior set

ii. Continuous behavior

A continuous behavior set typically comprises a stream of values on an effort scale, with symbols used shown in Fig. 6.13. The text shorthand used is "e".

6.3.5.2 Determined Behavior and Probable Behavior

Behavior elements within a discrete behavior set can be further categorized as "determined" or "probable" behavior based on the probability of behavior occurrence.

i. Determined behavior

The term "determined behavior" means that if the actor has selected the behavior in question the behavior will occur with 100 % probability. In other words, whether the behavior would occur or not rests completely on the actor with no other factors in play.

The symbol for "determined behavior" is as shown in Fig. 6.12.

ii. Probable behavior

The term "probable behavior" means that if the actor has selected the behavior in question the behavior may occur with a probability lesser than 1. This situation is typically caused by the lack of behavioral resources or opportunity. Within the behavior management institution, certain determined (and undesirable) behaviors can be changed to probable behaviors through a restriction of resources or opportunities for such behavior.

b_2 in Fig. 6.14 is a probable behavior, which is notated as "$p(b_2)$". The appropriate text notation is added in parentheses behind the probability symbol p.

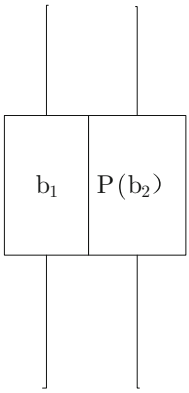

Fig. 6.14 b_2 is the binary behavior for the probable behavior

6.4 Numbering Rules for Behaviors and Components

6.4.1 Numbering Rules for Behaviors

6.4.1.1 How to Express a Discrete Behavior Set

For a discrete behavior set, if we are looking at the Sun Diagram of a single-actor institution, we can use subscripts to annotate various behaviors. Generally speaking, the more attention is to be paid to the behavior, the larger the number of this subscript. This is because in the Sun Diagram we generally deploy several components for a behavior that needs specific attention, which means that there would be more space to the right for the laying out of such components with key behaviors with larger numbers. For example, within the penalty institution the deployment of both the observer and suppressor is needed to tackle undesirable behavior. Therefore, b_1 stands for normal behavior, and b_2, undesirable behavior. In the bonus institution, the deployment of the observer and promoter is needed with regard to desirable behavior, and hence b_1 stands for normal behavior and b_2 indicates desirable behavior.

As for multiple-actor institutions, we add a dash ("–") behind the behavior number (the behavior set sequence number for the same actor) and then a number to indicate the actor in question. For instance, b_{1-2} indicates that the behavior in question is the first behavior for actor 2.

6.4.1.2 How to Express a Continuous Behavior Set

For a single-actor institution with a single behavior set shown in the Sun Diagram, the numbering for a continuous behavior set would not be needed.

In the case of a multiple-actor institution, we use the subscript on the symbol for effort levels to differentiate between the behavior sets for different actors. For example, e_1 indicates actor 1's effort level (i.e., the continuous behavior set of actor 1).

6.4.2 Numbering Rules for Institution Components

6.4.2.1 Institution Components That Are the Only One of Their Kind in a Single-Actor Institution Need Not Be Numbered

In a single-actor institution, if a certain component is the only one of its type in the Sun Diagram, then it does not need to be numbered. Just a text abbreviation that functions as the component label would do. For instance, r indicates the only promoter in the Sun Diagram, while s indicates the only suppressor.

6.4.2.2 When the Same Component Type Is Used for Multiple Behaviors, the Component Type and the Target Behavior Number Must Be Used in the Component Label

If, in the Sun Diagram there are multiple components of the same type for various behaviors, then the component should be labeled using the text abbreviation and the number of the behavior the component is targeted at. For instance, r_1 indicates that it is the promoter for behavior b_1 in a single-actor institution, while s_1 indicates that it is the suppressor for behavior b_1 in a single-actor institution. r_2 indicates that it is the promoter for behavior b_2 in a single-actor institution while s_2 indicates that it is the suppressor for behavior b_2 in a single-actor institution.

Here, we should note that if components of the same type are used for multiple behaviors, and these components are identical, these components can be labeled with the same number with no further differentiation needed. For example, the text symbol for the promoter for behavior b_1 is r_1, while the text symbol for the promoter for behavior b_2 is also r_1.

6.4.2.3 When Multiple Components of the Same Type Are Used for the Same Behavior, the Component Type, the Target Behavior Number and the Component Sequence Number Must Be Used in the Component Label

If multiple components of the same type are used for the same behavior, the component type, the target behavior number and the component sequence number must be used in the component label. For instance, r_{11} indicates that it is the first promoter for behavior b_1 in a single-actor institution, while r_{12} indicates that it is the second promoter for behavior b_1. On the other hand, r_{11-1} indicates that it is the first promoter for the first behavior b_{1-1} in a multiple-actor institution while s_{12-2} indicates that it is the second suppressor for behavior b_{1-2} in a multiple-actor institution.

6.4.2.4 The Actor Number in the Component Label May not Be Omitted in an Institution with Multiple Actors

We must note that with a multiple-actor institution each component label must also contain a specific actor reference. For example, even if in the institution a certain promoter, the promoter for the first behavior b_{1-2} of the second actor, is the only promoter deployed in the entire institution, we would still need to use r_{1-2} in the label and not omit the actor and behavior numbering as in r.

6.4.2.5 There Is the Possibility of Ambiguity in Certain Cases, But This Is Not Common

One more thing we would need to note is, with this expression method if the number of elements within an actor's behavior set is numerous, such as when it hits a double-digit count, confusion and ambiguity may occur. For instance, r_{11} can be understood both as the first promoter for b_1 (when there are multiple promoters used for b_1) and as the promoter for the 11th behavior b_{11} in a single-actor institution (when there is only one promoter used for the 11th behavior).

Fortunately, in the actual Sun Diagram, the aforementioned ambiguity essentially does not happen as there are basically no behavior sets with as many as 10 or more behavior elements within. Further, there are very few cases where 10 or more components of the same type are deployed with regard to the same behavior.

6.5 Drawing Rules for the Sun Diagram

6.5.1 Positioning Rules

In the Sun Diagram, the position of various line segments that connect various components to the behavior set is significant. For all components that have a promoter effect on behavior, lines are linked to the upper part of the symbol for the behavior while for components that have a suppressor effect, lines are linked to the bottom of the symbol for the said behavior. However, we should note that this rule is not absolute. For instance, an observer can be connected to the upper part of the behavior symbol and at the same time be connected to a promoter from one output terminal and to a suppressor from another output terminal.

6.5.2 Rules Regarding Component Terminals

Rules regarding component terminals include those concerning the positive/negative terminals as well as the positioning of terminals on various components. Components such as promoters and suppressors are represented

visually by a rectangle with rounded corners or an elliptical form. Both these forms have a long side and a short side. Here, we set the short side of the component as the "positive" terminal or end, and the long side of the component as the "negative" terminal or end. As such, both the symbols for the promoter and the suppressor have two positive terminals and two negative terminals.

Where the line that points towards the positive terminal origin is the factor that "produces" or "enhances" the effect of the said component, while the origin of the line that points towards the negative terminal is what "diminishes" the component effect (Fig. 6.15).

In Fig. 6.15, the behavior b_{1-1} is linked to the positive terminal of the promoter r_{1-1}, which means that r_{1-1} is produced as a result of the appearance of b_{1-1} or that it is enhanced due to the increase in effort level by b_{1-1}. The negative terminal of r_{1-1} is linked to behavior b_{2-2}, which means that r_{1-1} will decrease as the effort level of b_{2-2} rises.

In Fig. 6.15, the behavior b_{2-1} is linked to the positive terminal of the suppressor c_{2-1}, which means that c_{2-1} is produced as a result of the appearance of b_{2-1} or that it is enhanced due to the increase in effort level by b_{2-1}. The negative terminal of c_{2-1} is linked to behavior b_{2-2}, which means that c_{2-1} will decrease in power or efficacy as the effort level of b_{2-2} rises.

6.5.3 *Rules Regarding Line Segments (Directionality)*

The straight line without an arrow head indicates a two-way relationship, i.e., the components on either end of the line have an influence on each other. The straight line with an arrow indicates a one-way relationship where the origin component of the line has an impact on the component the line is pointing to.

In Fig. 6.16, the plain line that originates from e_1 points towards the positive end of promoter (reward) r_1, which means that as the effort level e_1 of actor 1 increases

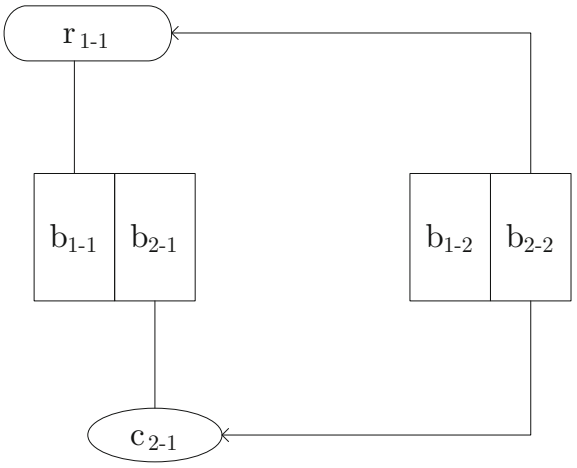

Fig. 6.15 The positive and negative ends of the component

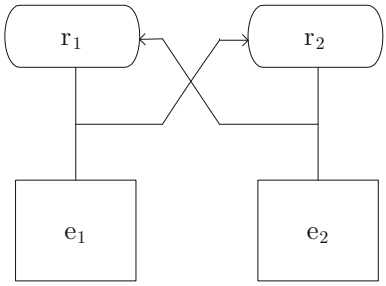

Fig. 6.16 Line with *no arrow*, and line with *arrow*

r_1 will also be augmented. Conversely, as r_1 increases e_1 will also be enhanced. The line segment indicates that there is a two-way relationship between r_1 and e_1. The arrow that originates from e_2 points towards the negative terminal of the promoter r_1, which means that as the effort level e_2 of actor 2 increases r_1 will decrease in value. The arrow indicates that while e_2 has an impact on r_1, r_1 does not have an impact on e_2. Similarly, the plain line that originates from e_2 points towards the positive end of promoter r_2, which means that as the effort level e_2 of actor 2 increases, r_2 will also be augmented. Conversely, as r_2 increases e_2 will also be enhanced. The line segment indicates that there is a two-way relationship between r_2 and e_2. The arrow that originates from e_1 points towards the negative terminal of the promoter r_2, which means that as the effort level e_1 of actor 1 increases, r_2 will decrease in value. The arrow indicates that while e_1 has an impact on r_2, r_2 does not have an impact on e_1.

6.5.4 Rules on the Omission of Behavioral Cost

There is a cost associated with every behavior or action. In contrast to behavioral reward, behavioral cost has already been put in at the point of behavior occurrence, and it is a necessary expense. Therefore, we do not need to express cost using a probability (or rather, the probability here is always 1 with cost). On the other hand, the reward is realized after the occurrence of the behavior at a certain probability. Therefore, in the Sun Diagram it must be expressed using that probability.

In the Sun Diagram, apart from special cases (such as when the behavior of other actors or when other components point towards the cost), when the cost is only related to the behavior at hand and not with other behaviors or components (i.e., it only has an impact on the behavior that has given rise to it or is impacted by such behavior), for the sake of simplicity we can omit the cost in the diagram. However, sometimes, we can also include the cost element in the diagram if we wish to emphasize the role that cost plays. However, we should note that when calculating the utility of behavior b_i, the equation should include the cost c_i of such behavior.

6.6 The Five Basic Institution Structures in the Sun Diagram

In Chaps. 7, 9 and 10 of this volume we will see that in terms of basic institution features, there are three fundamental institution types in use in human society: the behavior management institution, the task sharing institution, and the benefits distribution institution.

Of these three institution types, the function of the behavior management institution is to manage the behavior of individuals targeted for management, while the job of the task sharing institution is to assign tasks to multiple institution individuals. The objective of the benefits distribution institution is to divide "benefits" among multiple individuals.

When we use the Sun Diagram to analyze these three key institutions, we find that:

> In the behavior management institution category, in reality we find that two institution sub-types exist, i.e., the penalty institution and the bonus institution. The Sun Diagram for both institutions is similar in structure, which means that the penalty institution and the bonus institution are structurally the same kind of institution.

Within the task sharing institution category, there are only three distinct sub-types found using the Sun Diagram: the double-independent institution, the reward-sharing institution, and the cost-pooling institution. All other sub-types in this category are adaptations of, or variations on, these three key types.

Using the Sun Diagram, only one type of benefits distribution institution was found: the competitive institution, with all other sub-types in this category adaptations of, or variations on, this basic sub-type.

Hence, based on analysis using the Sun Diagram, we have found five basic institution structures in existence in human society. The complex institutions that we see in practice can be regarded as adaptations of, or variations on, these five basic institution structures.

These three types of institutions can be categorized based on institution function. Analyzing by way of the Sun Diagram, we find that these three key types all have distinctive institution features. An observer can always be found in the Sun Diagram for the behavior management institution, while in the task sharing institution, the common use of the reward and cost in the Sun Diagram for various basic institution structures is regularly arranged. In the Sun Diagram for the benefit distribution institution, promoter interactions are salient. We see that the Sun Diagram is indeed a scientific and powerful tool for the analysis of institution structures.

Below, we look briefly at the Sun Diagram characteristics of these five institution structures. Detailed analyses of these institutions can be found in Chaps. 7, 9 and 10 of this volume.

6.6.1 *The Basic Structure of the Behavior Management Institution*

The function of the behavior management institution is to encourage the individual to abandon undesirable behavior and to opt for desirable behavior. Examples include the "high pay for non-corrupt officials" institution, institutions designed to encourage hard work among students or workers.

Behavior management institutions are commonly divided into two broad types: the penalty institution and the bonus institution. In the Sun Diagram, we see that observers are deployed in both institutions. The observer in the penalty institution is targeted at undesirable behavior while the observer in the bonus institution is targeted at desirable behavior. From the perspective of the Sun Diagram, these two institutions have similar structures.

6.6.1.1 The Penalty Institution and the Bonus Institution Share a Similar Structure

The Sun Diagram for the penalty institution (with the penalty institution with the binary behavior choice as an example) is as shown in Fig. 6.17.

The Sun Diagram for the bonus institution (with the bonus institution with the binary behavior choice as an example) is as shown in Fig. 6.18.

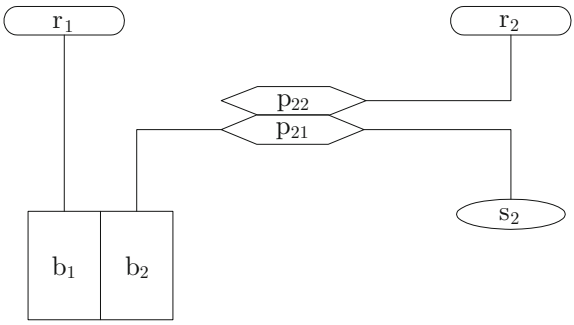

Fig. 6.17 Sun Diagram for penalty institution for binary behavioral choice

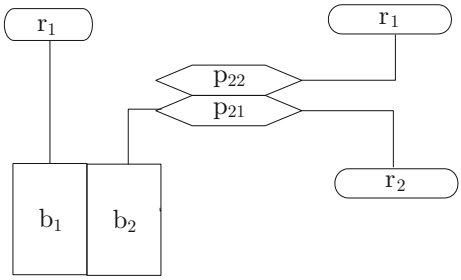

Fig. 6.18 Sun Diagram for bonus institution for binary behavioral choice

6.6.2 *The Three Basic Institution Structures for Task Sharing Institutions*

The task sharing institution works by parceling pieces of a larger task or job to multiple individuals and by encouraging them to complete the tasks assigned. One key issue that the task sharing institution must address is how to ensure that the individual puts in a reasonable amount of efforts in the target behavior so that tasks may be completed with efficiency and without over-consumption of resources.

From the perspective of the Sun Diagram, the task sharing institution has three basic structures: the double-independent structure, the reward-sharing structure, and the cost-pooling structure.

6.6.2.1 The Double-Independent Institution

The Sun Diagram for a double-independent institution is shown as in Fig. 6.19.

Within the double-independent institution, each individual within the group has exclusive enjoyment of his behavioral reward and at the same time bears his own behavioral cost entirely. The double-independent institution is mainly used to divide tasks between individuals, such as the unloading of goods and the contracting of farm work by lots.

6.6.2.2 The Reward-Sharing Institution

The Sun Diagram for a reward-sharing institution is shown as in Fig. 6.20.

The reward-sharing institution is an institution that comprises multiple individuals who work together as a collective to complete a larger task. The distinctive feature of this institution is that the reward from collective output is divided equally

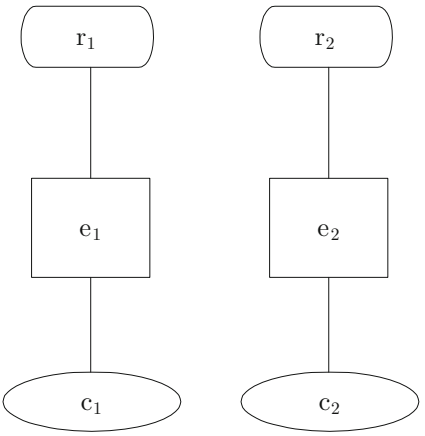

Fig. 6.19 Sun Diagram for double-independent institution (two individuals)

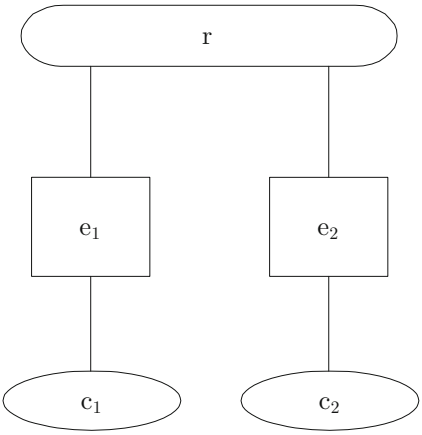

Fig. 6.20 Sun Diagram for reward-sharing institution (two individuals)

amongst all participants. However, each individual bears his behavioral cost separately and individually.

This type of institution is often used with production that generates positive externalities, or when total output cannot be divided among individual. Examples include the repair of public roads and bridges and environmental cleanup efforts.

6.6.2.3 The Cost-Pooling Institution

The Sun Diagram for a cost-pooling institution is shown as in Fig. 6.21.

The cost-pooling institution is also an institution where multiple individuals work together to complete a larger task. The distinctive feature of this institution is that while the cost of productive behavior is shared equally by all individuals, each individual has exclusive enjoyment of the reward granted to him. Very often,

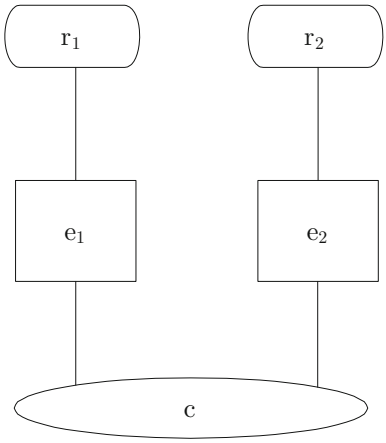

Fig. 6.21 Sun Diagram for a cost-pooling institution (two individuals)

behaviors that generate negative externalities are managed using a cost-pooling institution. For example, when the production of a certain good causes environmental pollution, while the benefits of such production go directly to the enterprise, the whole of society has to bear with the resulting pollution. In another example: while the operating income of a logistics company goes directly to the company, the traffic congestion caused by its transport behavior has to be borne by society in general.

The one similarity that the three sub-types of the task sharing institution share is that in terms of the Sun Diagram, the common use of the reward and cost is regularly arranged. With the double-independent institution, there is no link between various individuals. In other words, they are independent both in terms of behavioral cost and reward. On the other hand, with the reward-sharing institution, the reward is shared between various individuals and in the cost-pooling institution, the cost is shared between various individuals.

In Chap. 9, we will see that these differences in the Sun Diagram structure cause dramatic differences in the effect of these various institution structures. With the double-independent institution, individual effort level is optimized, while in the reward-sharing institution, individual effort level is comparatively lower, leading to low productivity and efficiency. On the other hand, in the cost-pooling institution, individual effort level tends to be comparatively high, leading to the over-consumption and depletion of resources.

6.6.3 Benefit Distribution Institutions: The Competitive Institution

Benefit distribution institutions are a kind of institution that seeks to give out "benefits" to each and every individual within the institution. In social life, many "benefits" are often limited and there is not enough to go around for everyone. Therefore, the benefits available must be distributed to individuals in accordance with certain standards or principles to maximize the effect of these benefits in society. Examples of "benefits" are the opportunities to access higher education, to access jobs and job promotions, to win a greater market share for the enterprise's products etc. From an institution management perspective, these are also distributed to individuals through benefit distribution institutions.

There is only one basic structure for benefit distribution institutions, which is that of the competitive institution. The Sun Diagram for the competitive institution is as shown in Fig. 6.22.

Competition is the targeted behavior aimed at seeking benefits, and through which individuals seek to eliminate the adverse effects posed by other individuals. In reality, there are many instances of competition, such as tender exercises for building projects, the *gaokao* college admissions institution, the arms race between nations etc.

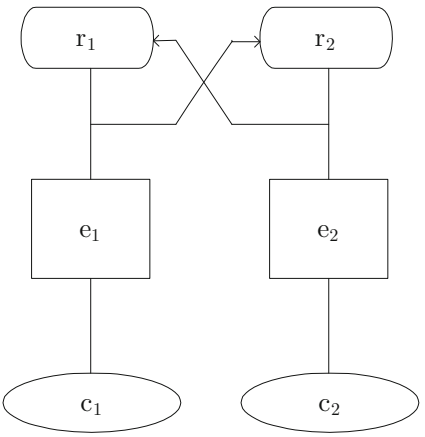

Fig. 6.22 Sun Diagram for competitive institution (two individuals)

The distinctive feature in the Sun Diagram for the benefits distribution institution is that, the interaction between individuals in terms of behavioral rewards is salient. For example, in the Sun Diagram for the competitive institution (Fig. 6.22), individuals place pressure on each other in terms of behavioral rewards.

6.7 Some Improvements and Variations on the Five Basic Institution Structures

In reality, to enhance institution effect, we often need to improve on the flaws of various basic institution structures. In addition, sometimes we would have to create new variations on basic institution structures to meet certain needs.

6.7.1 A Variation on the Binary or Two-Dimensional Penalty Institution: The Ternary Penalty Institution

In the penalty institution for binary or binary behavior choice, the actor's behavior set consists of two elements: the normal behavior and the undesirable behavior. However, in some circumstances there are behavior sets that contain three behavior elements. Take for example a penalty institution where employees are penalized for contravening workplace rules and regulations: employees can choose one of three possible behaviors, "adhere to the rules", "contravene the rules slightly", and "grave contravention of the rules". In this situation, we can then expand the two-dimensional penalty institution to create a ternary institution (Fig. 6.23).

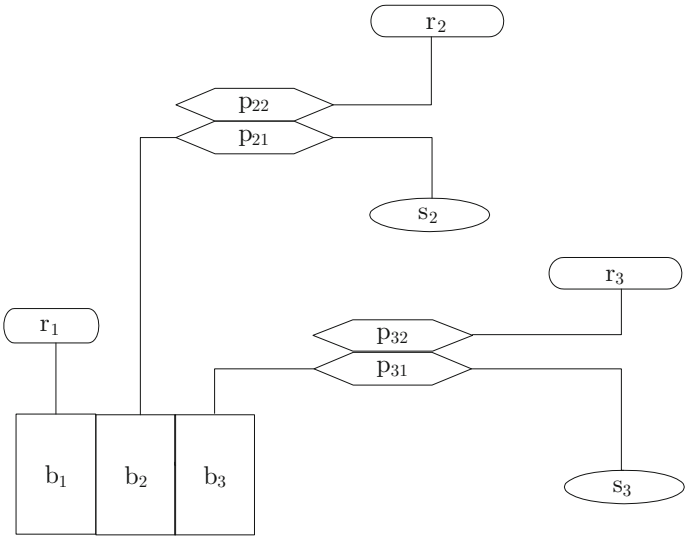

Fig. 6.23 Sun Diagram for penalty institution for ternary behavioral choice

6.7.2 *Improving the Reward-Sharing Institution: The Use of Cost Subsidies*

In Chap.9, the reader will see that compared to the double-independent institution, individuals in a reward-sharing institution tend to show a comparatively lower effort level. To address this issue, we can improve the reward-sharing institution with the use of cost subsidies.

The Sun Diagram for a reward-sharing institution with cost subsidies applied is shown as in Fig. 6.24. The reader will see that Fig. 6.24 is a variation of—and improvement on—Fig. 6.20 (i.e., the simple reward-sharing institution).

6.7.3 *A Variation on the Reward-Sharing Institution: The Parallel-Cooperation Institution*

Within the ordinary reward-sharing institution, group benefit is the sum total of individual benefits. However, in some cooperative groups, the total benefit from cooperation is greater than the sum total of individual benefit [should individuals go alone]. With these cooperative groups, the institution is correspondingly named the cooperation institution.

There is a kind of cooperation institution known as the "parallel-cooperation institution". (As to what the parallel-cooperation and serial-cooperation institutions

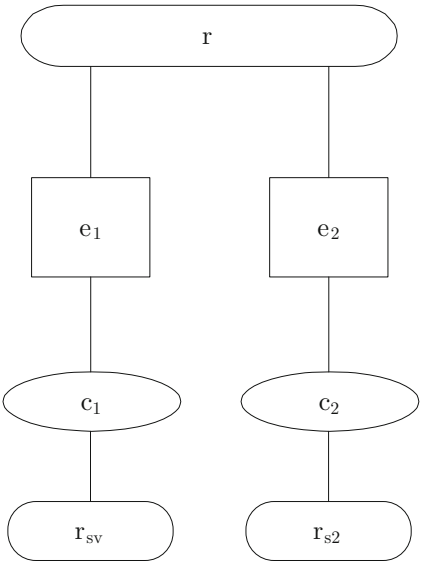

Fig. 6.24 The Sun Diagram for a reward-sharing institution with cost subsidies applied (two individuals)

are, please refer to Chap. 9 of this volume.) The Sun Diagram for such an institution is as shown in Fig. 6.25.

The reader will see that Fig. 6.25 is actually a variation on the Sun Diagram for the reward-sharing institution (Fig. 6.20), with the outcome "*con*" added to the original structure.

6.7.4 Improving the Cost-Pooling Institution: The Application of Tax Collections

The reader will see in Chap. 9 that compared to the double-independent institution, individuals in the cost-pooling institution tend to show a higher level of effort, which in turn can lead to the over-consumption of resources and thereby the lowering of production efficiency. To resolve this issue, we can make improvements to the cost-pooling institution with the use of tax collections.

The Sun Diagram for a cost-pooling institution with tax collections is as shown in Fig. 6.26. The reader will see that Fig. 6.26 is a variation of—and improvement on—Fig. 6.21 (i.e., the simple cost-pooling institution).

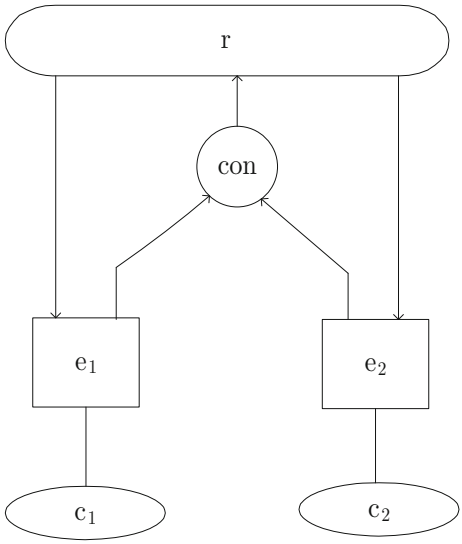

Fig. 6.25 Sun Diagram for parallel-cooperation institution (two individuals acting in concert)

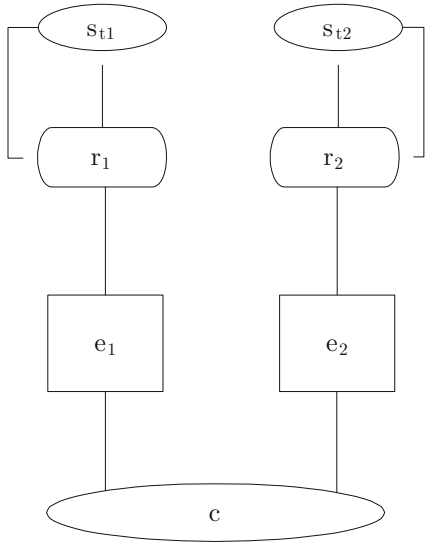

Fig. 6.26 Sun Diagram for a cost-pooling institution with tax collections (two individuals)

6.7.5 Improving the Competitive Institution: The Competitive Institution "with Cop"

In Chap. 10, the reader will find that in certain areas, a regular competition can become an adverse competition, causing losses to every party that engages in such

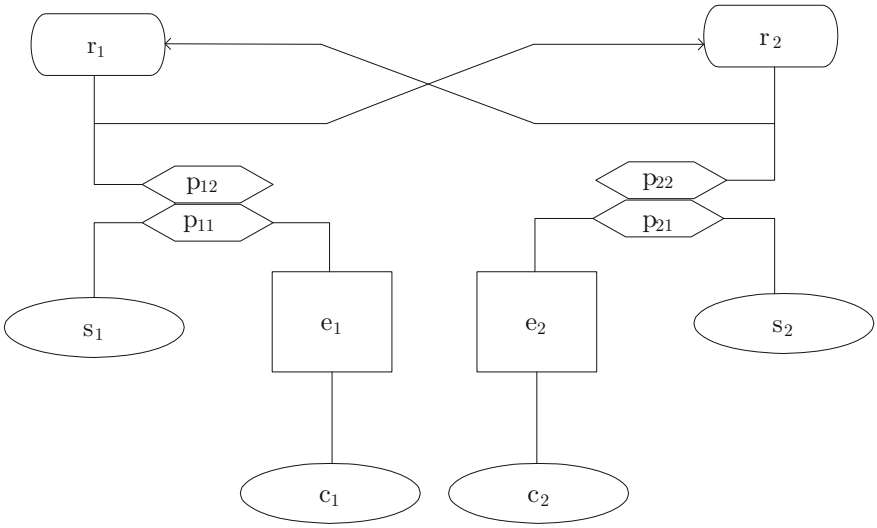

Fig. 6.27 Sun Diagram for competitive institution "with cop" (two individuals)

behavior. To resolve this issue, we can make improvements to the competitive institution with the use of the "cop function".

The Sun Diagram for a competitive institution "with cop" is as shown in Fig. 6.27. Although the diagram here is a little bit more complicated, the reader should still be able to see that Fig. 6.27 is a variation of—and improvement on—Fig. 6.22 (i.e., the simple competitive institution), with the additional deployment of an observer and suppressor on the effort level of both individuals in Fig. 6.27.

Chapter 7
Institutions of Behavior Management

In this chapter, we mainly study the impact of institution parameters and properties of institution components that affect institution performance, with attention paid to the effect of small changes made to various institutions. In fact, you would find that the structure of the Sun Diagram for the two behavior management institutions (the penalty institution and the bonus institution) to be examined in this chapter to be similar.

Behavior management institutions are chiefly used for controlling the behavioral choice of individuals so that they opt for behaviors that are aligned with management objectives.

Based on the various behaviors detected or observed, behavior management institutions can be categorized into two basic types: penalty institutions and bonus institutions. Of the two, penalty institutions observe for the occurrence of undesirable behavior while bonus institutions observe for the occurrence of desirable behavior.

The promoters and suppressors used in the penalty and bonus institutions are largely of the reward type. Moreover, some behavior management institutions make use of opportunity-type promoters or suppressors or deploy resource-type promoters or suppressors. However, if reward-type promoters or suppressors do not exist within these institutions they cannot be classified as penalty or bonus institutions. In reality, the use of penalty and bonus institutions are much more common to other types of behavior management institutions.

Finally, there is one more thing to be noted. Beginning from Chap. 7 of this book through to Chap. 10, when calculations of individuals' behavioral utility are conducted, we assume that these individuals (i.e., actors under management) are risk-neutral in their risk profiles. This way, when calculating behavioral utility, we can regard the difference between behavioral utility + behavioral reward and behavioral cost to have a positive linear relationship and hence use the difference between reward and cost (i.e., benefit or profit) as a direct expression of utility. This way, analyzing and calculating behavioral utility within an institution would be a far more straightforward process. At the same time, the data used for calculations are all fictional numbers used solely for the purpose of illustrating the calculation process and institution design principles.

S. Sun, *Five Basic Institution Structures and Institutional Economics*,
DOI 10.1007/978-981-10-0344-8_7

7.1 Penalty Institution of Binary Behavioral Choice

7.1.1 Basic Features of the Penalty Institution of Binary Behavioral Choice

The penalty institution seeks to eliminate or suppress undesirable behavior by individuals. Examples of penalty institutions include state criminal laws and law and order agencies, and punitive regulations for employees put in place by enterprises and public institutions together with the relevant implementing departments or personnel. The penalty institution is relatively common in human society and has a wide range of applications.

The distinctive feature of a binary or two-dimensional penalty institution is that there is only one individual actor under management and this individual has to choose between one of two possible behaviors within a binary behavior set. The two selections available under binary behavioral choice are: normal behavior b_1 and undesirable behavior b_2.

In the penalty institution, as long as the individual chooses normal behavior b_1, he would receive reward r_1. Hence, we see that in this institution there is necessarily a relationship between normal behavior b_1 and reward r_1. For example, as long as a hardworking student opts for taking an examination the right way (i.e., b_1), he or she would be guaranteed of a certain grade (i.e., r_1). As there is necessarily a relationship between normal behavior b_1 and reward r_1, normal behavior does not have to be b_1 monitored or observed under such an institution.

As the individual's possible behavior set is made up of two options, if the individual does not opt for normal behavior b_1, then he would have opted for the undesirable behavior b_2. However, as this institution does not observe for b_1, we cannot determine if b_1 has occurred (without loss of generality, the result r_1 of b_1 is sometimes not observable and sometimes observable. Therefore, we cannot absolutely be certain that r_1 is the result of b_1. For example, if a student were to take an examination and then pass it, we cannot assume that the student has passed the exam by "following all the exam regulations". The student could have either passed the exam the right way, or passed through "fraudulent means". Therefore, generally speaking, we cannot determine whether b_1 exists based on r_1, and cannot use b_1 to determine whether b_2 has occurred.

Therefore, we must monitor or observe for b_2. Therefore, a binary observer p_2 is deployed for the observation of undesirable behavior b_2. When undesirable behavior b_2 occurs (such as cheating during an exam), the observer p_2 is able to observe b_2 (i.e., the cheating behavior is discovered) at a probability of p_{21}. At this point, the individual would receive penalty s_2 (such as having his exam results voided) as a result of opting for behavior b_2. At the same time, as the performance of observer p_2 cannot be 100 % efficient (take for instance exam invigilation, which cannot uncover 100 % of cheating cases within the exam hall), the observer p_2 will malfunction (i.e., not uncover a case of cheating) at a probability of $p_{22} = 1 - p_{21}$, with b_2 not observed. At this time, as the observer has not detected the individual's

behavior b_2, the individual would still receive reward r_2 (such as in the case where a student passes the exam and obtains a certain grade as his cheating behavior was not uncovered).

7.1.2 Sun Diagram Within a Penalty Institution of Binary Behavioral Choice

The Sun Diagram within a penalty institution of binary behavioral choice is shown as in Fig. 7.1.

The fundamental feature of the Sun Diagram under the penalty institution of binary behavioral choice is that there is a binary behavior set that consists of normal behavior b_1 and undesirable behavior b_2. Normal behavior b_1 is directly linked to the promoter r_1. Undesirable behavior b_2 is directly linked to the binary observer p_2, with the "observation output (observed)" of the observer linked to suppressor s_2 while "observation output (not observed)" is linked to the promoter r_2.

7.1.3 Institution Parameters and Conditions for a Valid Institution

Parameters of a penalty institution of binary behavioral choice are shown in Table 7.1.

The condition for a valid binary behavior penalty institution is $u_1 > u_2$, i.e., $r_1 - c_1 > (1 - p_{21})r_2 + p_{21}s_2 - c_2$.

This gives us:

$$p_{21} > \frac{(r_2 - r_1) + (c_1 - c_2)}{r_2 - s_2} \tag{7.1}$$

That is to say, if observation intensity p_{21} is lesser than the requirements set out in Eq. (7.1), the institution is ineffective and the rational individual would opt for undesirable behavior b_2.

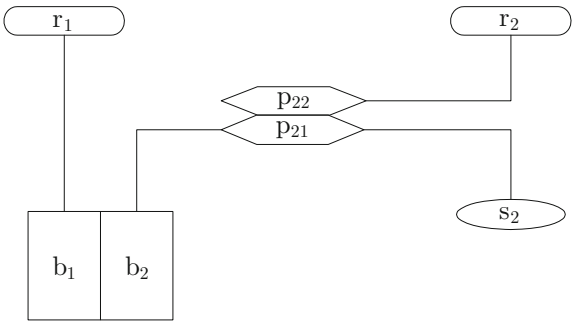

Fig. 7.1 Sun Diagram within a penalty institution of binary behavioral choice

Table 7.1 Parameters of a penalty institution of binary behavioral choice

Behavior	Content	Reward	Probability	Cost	Utility
b_1	Normal behavior	r_1	1	c_1	$u_1 = r_1 - c_1$
b_2	Undesired behavior	r_2	p_{22}	c_2	$u_2 = p_{22}r_2 + p_{21}s_2 - c_2$
		s_2	p_{21}		$= (1 - p_{21})r_2 + p_{21}s_2 - c_2$

If transforming the abovementioned equation, we get:

$$s_2 < r_2(1 - \frac{1}{p_{21}}) + \frac{r_1 + (c_2 - c_1)}{p_{21}} \tag{7.2}$$

Please pay attention to $(1 - \frac{1}{p_{21}}) < 0$ in the equation.

This means that with a specific observation intensity level p_{21}, penalties for undesirable behavior must meet the requirements set down in Eq. (7.2). Only under this circumstance, can the institution be effective. That is to say, it is only at this point that the rational individual would opt for normal behavior. (Please note: under normal circumstances $s_2 < 0$, and at this point the smaller s_2 is, the greater $|s_2|$ will be. That is to say, the smaller the value of s_2, the stronger the penalty.)

7.2 Typical Parameter Example for Penalty Institution of Binary Behavioral Choice, Part I: The Presumption of Innocence

7.2.1 About the Institution of the "Presumption of Innocence" and Its Widespread Use

In the area of jurisprudence, a common institution that is in use is the institution of "the presumption of innocence" (or "innocent till found guilty"). The principle was first proposed in the *Declaration of the Rights of Man and of the Citizen* of 1789 in France. Article 9 of the *Declaration* reads: "… all persons are held innocent until they shall have been declared guilty…" Currently, "the presumption of innocence" is a fundamental principle that underpins criminal law institutions in various countries. Article 162 of the China's *Criminal Procedure Law* also states that: "a judgment of innocence if the evidences submitted are so insufficient that the defendant cannot be found guilty of the crime accused, founded."

In practice, "the presumption of innocence" is not merely a jurisprudential principle, it is also a common behavior management institution. Take for example the fixed-wage institution for public servants, and the institution wherein an appointee is considered to have completed his term successfully if no major

incidents have occurred during his tenure. These are all examples of the individuals involved "presumed innocent till found guilty".

This type of institution is particularly useful when normal behavior b_1 is more challenging to observe compared to undesirable behavior b_2. For example, the work performance of a public servant can be hard to monitor and observe. Hence, his wage levels cannot be determined by 'workload'. The only thing his superiors can do is to observe for obvious instances of dereliction of duty. When such instances have not been observed (even if they have occurred) the public servant would be treated as having acted normally and will receive his wages per normal. In terms of the Sun Diagram that reflects the institution structure, this type of wage institution is similar to the "presumed innocent till found guilty" principle in jurisprudence.

7.2.2 *Characteristics of, and Sun Diagram for, the Institution of the "Presumption of Innocence"*

In an institution of the "presumption of innocence", b_1 is normal behavior and b_2, undesirable behavior (such as various illegal acts).

The "presumed innocent till found guilty" institution is a special instance of a penalty institution of binary behavioral choice, with parameters featuring $r_2 = r_1$. If undesirable behavior b_2 has not been observed, even if the individual has in reality opted for undesirable behavior b_2, he can only be treated as if he has opted for normal behavior b_1 with the reward for normal behavior $r_2 = r_1$ granted to him. If undesirable behavior has not been detected, the reward granted is similar to the reward granted for normal behavior. Of course, if behavior b_2 has been detected, then penalty s_2 is meted out to the individual.

In the Sun Diagram for "presumed innocent until found guilty" institution if the individual opts for undesirable behavior b_2 then the probability the behavior b_2 being detected is p_{21} while the probability of the behavior b_2 not being detected, i.e., recognized falsely as the individual having opted for the other behavior b_1, is $p_{22} = 1 - p_{21}$ (Fig. 7.2).

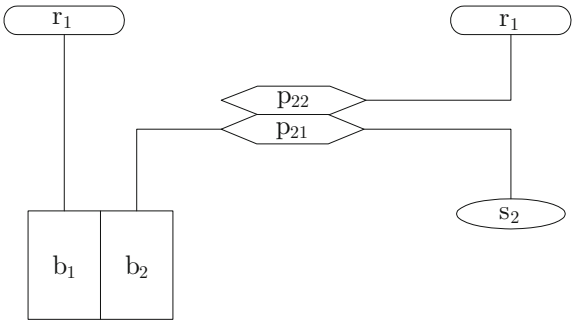

Fig. 7.2 Sun Diagram for the "presumption of innocence" institution

7.2.3 Institution Parameters and Conditions for a Valid Institution of the "Presumption of Innocence"

7.2.3.1 Institution Parameters

Institution parameters for the institution of the "presumption of innocence" are as shown in Table 7.2.

7.2.3.2 Conditions for a Valid Institution

If the "presumed innocent till found guilty" institution is effective, then the rational individual under management would opt for normal behavior b_1 and not for undesirable behavior b_2. Therefore, $u_1 > u_2$ would arise, i.e., $r_1 - c_1 > (1 - p_{21})r_1 + p_{21}s_2 - c_2$.

(i) Minimum level of observation intensity

The above equation then gives us:

$$p_{21} > \frac{c_1 - c_2}{r_1 - s_2} \tag{7.3}$$

That is to say, if we want to ensure that the institution works, that is, if we want the individual to opt for normal behavior b_1 under the institution, the level of observation intensity p_{21} must meet requirements set down in Eq. (7.3).

(ii) Requirements for the difference $r_1 - s_{21}$ between income from normal behavior and penalty for undesirable behavior

Transforming the equation above, we get:

$$r_1 - s_2 > \frac{c_1 - c_2}{p_{21}} \tag{7.4}$$

This means that with a specific observation intensity level p_{21}, the difference between the reward for normal behavior and penalty for undesirable behavior $(r_1 - s_2)$ must meet the requirements set down in Eq. (7.4). Otherwise, the institution will become non-effective. That is to say, it is only at this point that the rational individual would opt for normal behavior.

Table 7.2 Institution parameters for the institution of the "presumption of innocence"

Behavior	Content	Reward	Probability	Cost	Utility
b_1	Normal behavior	r_1	1	c_1	$u_1 = r_1 - c_1$
b_2	Undesired behavior	r_1	p_{22}	c_2	$u_2 = p_{22}r_1 + p_{21}s_2 - c_2$
		s_2	p_{21}		$= (1 - p_{21})r_1 + p_{21}s_2 - c_2$

7.3 Typical Parameter Example for Penalty Institution of Binary Behavioral Choice, Part II: Penalty Institutions Designed to Tackle Corruption

7.3.1 The Concept of Corruption (Corrupt Behavior)

Corruption refers to the abuse of one's official power to gain an unfair advantage or undue benefits. Corruption is common in countries around the world. Not only does this phenomenon cause economic loss to nations, it also leads to social injustice and has a negative impact on social stability. Hence, corruption also poses a very real threat to the nation's long-term stability and security. Due to the possible harms of corrupt behavior, many countries have implemented institutions designed to address this issue, such as mandatory information disclosure, administrative hearings, the reporting of personal finances, making full use of the "watchdog" function of the media, etc.

7.3.2 Characteristics of, and Sun Diagram for, Penalty Institution Designed to Tackle Corruption

In penalty institutions designed to address the problem of corruption, the individual's behavior set is also two-dimensional, with b_1 being normal behavior and b_2 being corrupt behavior, such as the acceptance of bribes. The task of institution design is to try and induce the individual to opt for the normal behavior within the behavior set and move away from the selection of corrupt behavior.

The penalty institution designed to address the problem of corruption is a special type of penalty institution with binary behavior choice. When the utility of corrupt behavior is greater than the utility of normal behavior, institution parameters feature $r_2 - c_2 > r_1 - c_1$. Here, we can regard r_2 as comprising "normal income + bribery income". As r_1 is normal income, it is not difficult to see that income from corruption r_2 will certainly be greater than normal income r_1, while the difference in behavioral cost between both options, c_2 and c_1, is comparatively small. In a

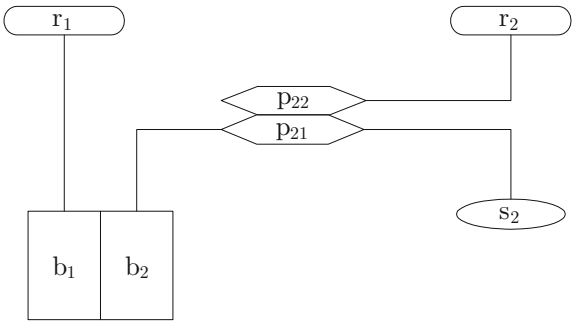

Fig. 7.3 Sun Diagram for institution designed to tackle corruption

situation where there is no suppressor for corrupt behavior (i.e., when no observers p_2 and penalties s_2 exist), $r_2 - c_2 > r_1 - c_1$ will always be true. At this juncture, the rational individual will certainly opt for corrupt behavior b_2 (Fig. 7.3).

7.3.3 Institution Parameters and Conditions for Penalty Institution Designed to Tackle Corruption

Institution Parameters for a Penalty Institution Designed to Tackle Corruption are shown in Table 7.3.

7.3.3.1 Institution Conditions for Penalty Institution Designed to Tackle Corruption

When $u_1 > u_2$:

$$r_1 - c_1 > (1 - p_{21})r_2 + p_{21}s_2 - c_2 \tag{7.5}$$

The actor will opt for normal behavior rather than engage in corruption. Therefore, the above equation shows the institution conditions required for the effective operation of a penalty institution designed to tackle corruption.

7.3.4 Institution Parameter Principles for Various Policy Measures Against Corruption in Management Practice

7.3.4.1 Minimum Normal Income r_1 and "High Salary for Clean Government"

Transforming Eq. (7.5), we get:

$$r_1 > (1 - p_{21})r_2 + p_{21}s_2 + (c_1 - c_2) \tag{7.6}$$

Table 7.3 Institution parameters for penalty institution designed to tackle corruption

Behavior	Content	Reward	Probability	Cost	Utility
b_1	Normal behavior	r_1	1	c_1	$u_1 = r_1 - c_1$
b_2	Undesired behavior	r_2	p_{22}	c_2	$u_2 = p_{22}r_2 + p_{21}s_2 - c_2$
		s_2	p_{21}		$= (1 - p_{21})r_2 + p_2s_2 - c_2$

Equation (7.6) shows that if we want the institution to work against corruption, income from normal work r_1 must be higher than the value to its right. That is to say, if normal income levels are too low, individuals would be more prone to engaging in corruption. This provides theoretical support for the "high pay for zero corruption" policy certain governments have adopted with their public servants.

7.3.4.2 Minimum Penalty Level s_2 and the "Severe Penalties for Corruption" Policy

Transforming Eq. (7.5), we get:

$$s_2 < r_2 - \frac{(r_2 - r_1) + (c_1 - c_2)}{p_{21}} \tag{7.7}$$

If we are to consider penalty level s_2, then we must have levels that satisfy Eq. (7.7) for the institution to be effective against corruption. Here, readers should note that in reality, the greater the value of punitive reward s_2 (note that this refers to numerical value, i.e., the further to the right the number is on the number axis, the greater the value), the lighter the punishment, while the small the number (the further to the left the number is on the number axis), the heavier the penalty. For instance, the value of penalties such as fines and job demotions are negative.

Equation (7.7) shows that if we want the institution to be effective against corruption we must have penalty levels large enough to make the value of s_2 smaller than the value to its right (note that under normal circumstances $s_2 < 0$, at this time the larger $|s_2|$ is, the smaller s_2 will be and the smaller the value of s_2, the stronger the penalty.) That is to say, if penalty levels are too low, individuals would be more prone to engaging in corruption. This provides theoretical support for the "severe penalties for corruption" policy certain governments have adopted with their public servants.

7.3.4.3 Minimum Level of Observation Intensity p_{21} and "Information Disclosure"

From Eq. (7.5), we get:

$$p_{21} > \frac{(r_2 - r_1) + (c_1 - c_2)}{r_2 - s_2} \tag{7.8}$$

That is to say, the observation rate must satisfy conditions in Eq. (7.8) for the institution to be effective. This provides certain theoretical ground to the "informational disclosure" policy adopted in several countries in their respective fights against corruption such as "transparent remuneration", "open enrollment", and "open recruitment".

From Eq. (7.8), we can also see the restrictive relationship between observation intensity p_{21} and other institution parameters:

The more income from corruption r_2 exceeds income from normal behavior r_1, the greater $(r_2 - r_1)$ is. This means that observation intensity p_{21} must also be enhanced to maintain institution effectiveness.

When income from corruption r_2 remains unchanged (under normal circumstances $r_2 > 0$), the greater the penalty s_2 (we should note $s_2 < 0$, with the greater s_2 is, the greater the absolute value of s_2), i.e., the greater $(r_2 - s_2)$ is, observation intensity p_{21} can be decreased a little and the institution can continue to be effective.

7.3.4.4 The Restrictive Relationship Between Corruption Income r_2 and Observation Intensity p_{21}

Let us now look at the restrictive relationship between income from corruption (illegal income from corrupt behavior and which has not been discovered) r_2 and observation intensity p_{21} when all institution conditions needed for institution effectiveness are in place.

Seeking the partial derivative of the foresaid Eq. (7.8), we get:

$$\frac{\partial p_{21}}{\partial r_2} > \frac{(r_2 - s_2) - (r_2 - r_1) - (c_1 - c_2)}{(r_2 - s_2)^2} = \frac{(r_1 - s_2) - (c_1 - c_2)}{(r_2 - s_2)^2} \tag{7.9}$$

From the above equation, we can see that when normal income r_1 is sufficiently large and when penalties are adequately harsh, i.e., when $(r_1 - s_2)$ is sufficiently large, $(r_1 - s_2) - (c_1 - c_2) > 0$ will arise and lead to $\frac{\partial p_{21}}{\partial r_2} > 0$. That is to say, an originally effective institution against corruption can only remain effective if observation intensity p_{21} increases in tandem with an increase in income from corrupt behavior r_2.

7.3.5 Case Study 1: The Institutional Root of Corruption in the Qing Imperial Household Department[1]

A popular folk ditty during the Qing Dynasty went: "Sparkling house, young trees, new paintings/Here, man of the Imperial Household Department is dwelling'." This referred to the fact that during that time, a member of the agency was very likely to become a rich man overnight (hence the sparkling new home, newly planted trees and 'modern' paintings in the home). It was said that agency officials "enriched themselves and hatched fraudulent schemes as if such acts were mandated [by the

[1]This case study was excerpted from Zhang (2010, May 15). Narrative edited for use in this volume.

sages]", and were so flagrantly corrupt that they had only stopped short of having treasury gold shipped directly to their homes (Zhang 2010).

7.3.5.1 About the Qing Imperial Household Department

The Imperial Household Department was an imperial agency set up between the reigns of Emperors Shunzhi and Kangxi in the Qing Dynasty. The agency was run by what were known as *baoyi*, or bondservants of the imperial family. One area of responsibility for the Imperial Household Department was the management of palace eunuchs and maids. Therefore, eunuchs were prone to siding with the agency at times of conflict. Due to the heavy responsibilities given to the Department, a strict management institution was drawn up for the agency.

For example, the Department was also responsible for the safekeeping of the Emperor's treasures, such as gold, silver, pearls, jade, coral, agate objects, etc. As such, the Emperor ordered that 25 men guard the Department treasury day and night, with the keys to the treasury held not by the guards but the guards of Qianqing Palace (the Emperor's audience hall). Multiple keys were also needed to gain access to the Treasury. Any time the treasury had to be opened, a number of officials were required to be present. Anyone who entered the treasury was subject to body searches upon exit. At the shutting of the treasury, a seal signed and stamped by multiple officials was applied on the treasury doors.

Each month, the Department was to create a written report of the income and expenses for the month, with the report accessible to the emperor at any time. Department officials were also required to report to the emperor in person at the end of the year.

Clearly, stringent rules had been drawn up for Department operations. Theoretically speaking, the institution would mean that it would be hard for agency members to pilfer even a single coin for themselves.

However, strangely enough, the more rules that were enacted, the more severe the problem of corruption became within the Department. It was as if a river dike had broken, with all awash in the waters of malfeasance. So what had happened exactly?

Although the management rules laid down for the Department were extremely strict, institution implementers were weak and the performance of institution components were poor. As a result, the management institution was rendered ineffective.

7.3.5.2 Other Cases

(i) The time the emperor wanted to have his pants mended

One day, Emperor Daoguang (who was known to be frugal) discovered a small hole in his pants where his knee had rubbed through. He ordered the Imperial Household

Department to have the piece of clothing patched. The pants were patched and then returned to the Emperor. When asked by the Emperor how much the repair had cost, the agency responded: 3000 taels of silver. (1 tael is about 40 g in today's measurements.) Daoguang was taken aback. Why did a simple patching job cost more than the tailoring of an imperial robe? This was the explanation offered by a member of the Department: "Your Majesty, your pants are made of a special brocaded silk. We had to cut patches out of several hundred bales of silk before we found an exact match for your pants. This is why the mending had cost more." If it had been a simple patching job, it would have cost five taels at the most, they said.

(ii) Curtains for the Emperor Guangxu

During preparations for Emperor Guangxu's wedding, a few embroidery workers were summoned to the palace to help with the sewing of curtains for the bridal chamber. The workers sat down to work after collecting bales of silk from the Imperial Household Department. The market price of this work was around 50 taels.

When the curtains were completed and readied for handover, the workers were 'urged' by the Department to quote a higher price for their services. And so they tried their luck: 500 taels, they said. The quote was turned back. *Try again*, the embroidery workers were told, *and this time higher*. And so they turned in an invoice for 1000 taels. The document went up to a Department official, who took one look and sent the invoice back. *Try again*, the workers were again told. The leader of the group decided to be bolder this round and submit an invoice for 5000 taels for the job. The other embroidery workers disagreed, fearing that the invoice would be rejected. Never did they know, when the invoice was reviewed again, the Department official in charge shook his head, and added another 20,000 taels to the invoice total in his writing. And this was how the Imperial Household Department purchased a set of curtains for the Emperor for the princely sum of 25,000 taels. Then, the Department retained 20,000 taels, giving the remaining 5000 taels to the embroidery workers who were too pleased with their windfall to sleep for days!

(iii) Bamboo scaffolding that cost 40,000 taels of silver

The best thing to happen to a member of the Imperial Household Department was not a job promotion, but a job assignment that would put him in charge of a new project. This is because a new project made for a great opportunity for agency staff to make a small (illicit) fortune. Once, during the late Qing period, the Emperor had wanted a bamboo scaffold erected inside the Forbidden City. And the final price tag on the scaffold? A shocking 40,000 taels of silver!

7.3.5.3 Institutional Reason 1: Observer Failure

One of the reasons for the rampant corruption in the Imperial Household Department was the fact that institution observers were essentially non-functioning, leading to low observation intensity p_{21}.

After learning that the Emperor Daoguang had ordered his clothes mended, in order to emulate the emperor, many officials took to tearing their clothes, and then having them patched subsequently. Once, Daoguang was struck by what military official Cao Wenzheng had worn to court, the clothing that had very clearly been mended. And so he asked Cao: "Tell me, how much did mending your clothes cost you?" Cao was about to give the emperor the answer—"a *qian* (5 g) of silver for ten holes"—when he noticed that the eunuchs present were staring at him with great nervousness. Beads of cold sweat began to form on Cao's forehead. He realized that the eunuchs must have inflated their claims for the same service, and so he said: "The price for mending a hole outside is three *qian* of silver." For information, three *qian* of silver at that time were enough to buy one a whole set of clothing. Cao thought that his quote would be pretty much in line with what Daoguang had paid for mending. Little did he expect Daoguang to sigh at his statement, saying: "It is cheaper outside after all! I had to pay five taels of silver to have a tear mended!" Cao staggered at the amount the emperor had mentioned.

The Emperor Guangxu was very fond of having eggs in his meals. At that time, an egg cost three or four copper coins on the market. However, the procurement price the Imperial Household Department reported was 30 taels of silver per egg.

Once, the emperor had asked the imperial tutor Weng Tonghe: "Sir, while eggs are truly tasty, they are also terribly expensive. Can you afford them, sir?" Weng replied slyly, "I only eat one or two eggs when we have rituals, and I do not dare purchase them otherwise." Because of this, Guangxu believed that eggs were a highly-priced commodity until the day he breathed his last. For each year, he had spent more than 10,000 taels of silver on the delicacy alone.

7.3.5.4 Institutional Reason 2: Strong Performance of the Corruption Promoter

The second key reason for the non-performance of the institution was that the performance of the promoter of corrupt behavior was just too strong, i.e., the level of benefits to be gained from corruption, i.e., r_2 was simply too high.

Patching a hole in the pants for 3000 taels of silver; making a set of curtains for the bridal chamber at the price of 25,000 taels; building a bamboo scaffold for 40,000 taels: the great part of these astronomical sums went into the lining of Department officials' pockets. Certainly, the large "profits" possible made for a key reason for the flagrant corruption that went on.

7.3.5.5 Institution Reason 3: Malfunctioning of Corruption Suppressor

Another reason for the rampant corruption among agency officials was the malfunctioning of the suppressor for undesirable behavior. This malfunction meant that penalty s_2 for undesirable behavior was severely lacking in degree.

The story of Emperor Guangxu's failed attempt to deal with a middle-level official in the Department illustrates this problem.

It was just before the occasion of Empress Dowager Cixi's birthday, and Guangxu had wanted to have a pair of bracelets made as her birthday present. And so he ordered the Department to make four prototypes. Cixi was extremely pleased at what she saw. When asked by Guangxu how much the bracelets would cost, the official replied: "Your Majesty, it is not expensive at all. Just 40,000 taels of silver! Guangxu was taken aback by the exorbitant quote. It just so happened that 40,000 taels made up the entirety of Guangxu's private savings.

Guangxu became suspicious of the Department official in charge of the matter. Following investigations, it emerged that the official responsible was indeed a very corrupt man who was no stranger to cooking the books. For the occasion of Cixi's 60th birthday celebration, the official had managed to take charge of all the procurement involved and made himself a tidy sum in the process.

And so Guangxu ordered the man to be stripped of his title and property on the charge of "contravening the institution".

Sometime later, the vacancy for the Salt Commissioner for Jiangxi Province came up. Strangely enough, the disgraced official was unanimously nominated for the position by all court officials involved, and thus became a fourth-grade official with substantial power (over salt revenues) in an instant. Guangxu had no idea what had hit him.

7.3.5.6 Institutional Reason 4: Loss of Control over Behavior Set (Control not Realizable with Institution Re-design)

Loss of control over the behavior set was a key reason why the Imperial Household Department was able to act without restraint for a long time. In institution design, by controlling the availability of behavioral resources or behavioral opportunity, in general we will be able to impact the elements within the individual's behavior set, i.e., the number and types of behaviors b_i in order to realize management of individual behavior. However, as the institution manager, the emperor was unable to 'design' the behavior elements within the institution behavioral set b_i, hence failing to revamp the institution.

Emperor Daoguang was most unhappy with the profligacy of the Imperial Household Department, and hence tried to tackle the problem of corruption within the agency.

He was very fond of sweet potato vermicelli in soup, and once ordered the agency to bring him a bowl of the snack. He waited for days on end for the dish, and finally summoned a member of the Department to find out what had happened to his food. *Your Majesty*, he was told, *we are still raising funds for the dish.* Daoguang flew into a rage when he heard this. *Just how much did a bowl of sweet potato vermicelli in soup cost?*

The official responded: "To make the dish, the imperial kitchen would first have to set up a dedicated 'Vermicelli Section', which meant that we needed more

manpower to write up the relevant regulations. Then, we would also have to procure ingredients for the dish. If you want to have vermicelli in soup, your Majesty, we would need an additional 60,000 taels of silver in the imperial kitchen budget."

Daoguang laughed coldly at the explanation, and replied: "There is no need to go through all that. Before I became emperor I had a bowl of sweet potato vermicelli at Qianmen Street for only two copper coins. Just arrange for someone to buy me a bowl from there every day."

A few days later, and the emperor still had yet to see his snack. Once again, he summoned the officials responsible and started to berate them. In response, he was told: "Your Majesty, we have gone to Qianmen Street, but we were unable to find hawkers selling sweet potato vermicelli in soup. We would travel further to look for other hawkers, except we fear that the soup would have gone cold by the time we come back. We are studying the problem right now!"

And what was the true story behind this? The Department had secretly sent men to chase away all the vermicelli sellers on Qianmen Street.

All Daoguang could do was to let out a sigh and say, *forget about the vermicelli, then.*

This tale tells us that as the "institution manager", the emperor was completely unable to control the behavior set of the Imperial Household Department officials. The emperor had known the street price of the dish, which meant that there was no way for Department officials to line their pockets had they heeded his order to purchase some outside the palace. This was why the Department did not agree that the emperor's proposal was a natural one. The problem was so severe that the agency was able to thwart the emperor's plans!

7.3.5.7 The Fundamental Reason for Institution Malfunction: Poor Implementer Performance

The fundamental reason for the poor performances of the institution components mentioned above, i.e., of the observer, the suppressor and (control over) the behavior set, etc., is the lack of usable managerial resources for the manager, i.e., weak institution implementers. We can just see how stacked the situation was against the emperor: one man at the top versus thousands at the agency who worked to cheat emperor in all sorts of ways. What this example shows is the result of institution component failure due to poor implementer performance.

7.3.6 Case Study 2: The Impact of Improving Observer Performance

7.3.6.1 Corrupt Behavior in a Coal Inspection Station

Typically, in China, a coal inspection station is set up at the exit from a coal-producing area. The function of the station is to weigh coal loads being

shipped from the district and to collect the corresponding tolls for these loads. Station staff was also responsible for verifying various tickets carried by truck drivers.

Investigations showed certain management stations to be severely plagued by corruption. In some cases, trucks were recorded by workers (who had accepted bribes) to be carrying less than what was in the actual load. This practice benefited both the station staff and the truck operator while massive losses to the state were caused at the same time. Over time, station staff had begun to itch for more benefits. Hence they began to allow trucks to detour round the coal inspection station entirely without having to pay the requisite tolls. As the losses to the state increased, so did the amount of bribes paid to station staff.

To cover their tracks, coal station staff members do not conduct such transactions in person. Instead, "cowboys" were hired to hang around the stations and to collect bribes from coal trucks taking the detour before splitting the takings with coal station staff. At the same time, coal station staff would try ways and means to force coal trucks to take the illegal detour (where bribes would be collected for passage). For instance, station staff might conduct truck inspections extremely slowly to hold up the line. It was not unheard of for trucks to wait in line for a whole day for the station inspection. At this point, the "cowboy" would "remind" truck drivers in the line that all they had to do was to take a certain detour for a quicker route at lower cost. As a result, a great many trucks ended up traveling by the illicit route.

Observations showed that as many as ten trucks took the illicit "shortcut" every hour, and based on fee collection standards of RMB1800 per truck this meant that at least RMB300,000 of bribes exchanged hands on a daily basis. As such, corruption at a single coal inspection station can cost the state over RMB100 million a year.

7.3.6.2 An Institution that Tackles Corruption at Coal Inspection Stations

(i) Sun Diagram for institution designed to tackle corruption

Based on the analysis above, the Sun Diagram for an institution designed to tackle corruption at a coal inspection station is as follows (Fig. 7.4).

(ii) Parameters for institution designed to tackle corruption at coal inspection stations

Current institution parameters are shown in Table 7.4.

From Table 7.4, we can see that in terms of various behavior options for station staff, "sending trucks on diversions" presents the greatest utility with "allowing

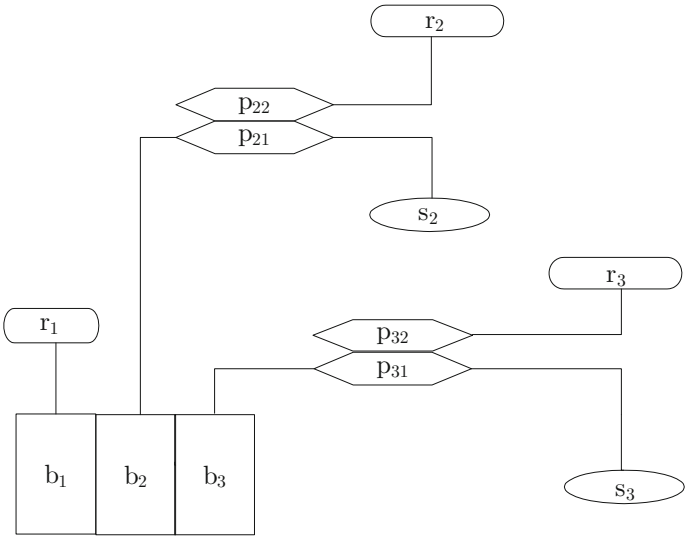

Fig. 7.4 Sun Diagram for an institution designed to tackle corruption at a coal inspection station (ternary behavior)

'lighter trucks' to pass through" in second place. The utility presented by "normal behavior" presents the least utility.

Therefore, the institution is ineffective.

(iii) Improvements to institution designed to tackle corruption at coal inspection stations

First, an electronic weighing apparatus shall be installed at the station. The apparatus automatically takes and displays the weight of the truck; weight data is then used by another automatic device to calculate fees payable, the amount of which is printed on a slip. This way, the weights of these trucks are no longer manually recorded by station workers. Hence, the workers have lost the opportunity to revise truck weights. As for the behavior that is "sending trucks on diversions", the opportunity to do so no longer exists.

Second, to address the issue of bribery, an electronic camera institution will be installed to enhance observer performance. The observation intensity vis-a-vis undesirable behavior was greatly improved as a result. With these improvements, the new coal station management institution is now a binary or binary behavior management institution, with the Sun Diagram for the institution shown in Fig. 7.5.

(iv) Parameters for institution designed to tackle corruption at coal inspection stations (post-institution improvement)

Current institution parameters are as shown in Table 7.5.

Table 7.4 Current institution parameters for an institution designed to tackle corruption at a coal inspection station

Component name	Component type	Element	Numerical value or characteristic	Description
Behavior set	Three-dimension behavior set	b_1	Normal coal management behavior	
		b_2	Allowing "lighter trucks" to pass through	
		b_3	Truck diversions	
Promoter of b_1	Reward	r_1	RMB200,000/year * 5 years = RMB1,000,000	Assuming the station employs four workers, with annual wage per person at RMB50,000 and assuming a work period of five years
Observer of b_2	Whistleblowing-type observer	p_{21}	0.1	Whistleblowing with no reward
		p_{22}	0.9	
Suppressor of b_2	Reward-type suppressor	s_2	RMB-200,000/year * 5 years = RMB-1,000,000	Dismissal, assuming the station employs four workers, with annual wage per person at RMB50,000 and assuming they could have worked another five years
Promoter of b_2	Reward-type promoter	r_2	RMB200,000/year * 5 + RMB2 million * 5 = RMB11 million	Normal wage + bribes from allowing "lighter trucks" to pass through, with work period at five years
Observer of b_3	Whistleblowing-type observer	p_{31}	0.1	Whistleblowing with no reward
		p_{32}	0.9	
Suppressor of b_3	Reward-type suppressor	s_3	RMB-200,000/year * 5 years + (RMB-1.2 million/year * 5 years) = RMB-7 million	Dismissal + jail term; assuming that the station employs four workers, with the loss to each staff member from dismissal standing at RMB-50,000/year and RMB-200,000/year for four members. Assuming that they could have worked another five years, total wage loss stands at RMB-1 million. The exchange utility for one person in jail is RMB-300,000/year and RMB-1.2 million/year for four persons. For a five-year jail term, total exchange utility stands at RMB-6 million

(continued)

Table 7.4 (continued)

Component name	Component type	Element	Numerical value or characteristic	Description
Promoter of b_3	Reward-type promoter	r_3	RMB200,000/year * 5 years + RMB100 million/year * 5 years = RMB501 million/year	Normal wages + bribes from letting "lighter trucks" pass, assuming such behavior can go on for five years
Utility of b_1		u_1	$u_1 = 100$	Assume that the individual under management is risk-neutral
Utility of b_2		u_2	$u_2 = 0.1 \times (-100) + 0.9 \times 1100 = 980$	Assume that the individual under management is risk-neutral
Utility of b_3		u_3	$u_3 = 0.1 \times (-700) + 0.9 \times 50100 = 45020$	Assume that the individual under management is risk-neutral

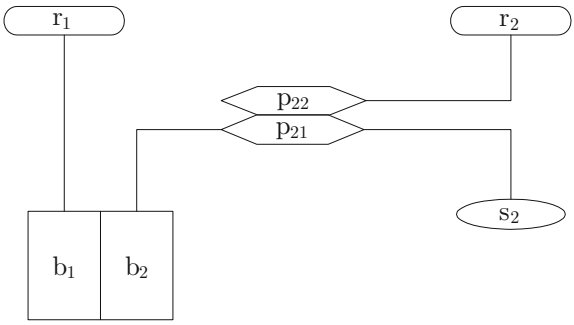

Fig. 7.5 Sun Diagram for an institution designed to tackle corruption at a coal inspection station (binary behavior)

From Table 7.5 we see that post-institution improvement, the utility of corrupt behavior has become far less than the utility of normal behavior. Under such an institution, the rational individual would not opt for corrupt behavior. This means that the institution is effective.

7.4 Typical Parameter Example for Penalty Institution of Binary Behavioral Choice, Part III: Penalty Institutions Designed to Tackle "Loafing on the Job"

7.4.1 Loafing

"Loafing on the job" in this context occurs when a member, or members of a group, present an effort level lower than that of others (this means that behavioral cost is also lower) but tries to obtain the same level of reward as the others. In modern society, laziness is commonplace as many social activities—such as labor—as conducted in a mass or communal manner. For example, in communal labor, there have been instances of individuals who would appear at the workplace reluctant to do any actual work, but still claim the same wage as others (who have actually worked) anyway.

7.4.2 Characteristics of Penalty Institution Designed to Tackle Problem of Loafing on the Job

Under such an institution, the behavior of "loafing on the job" is one of the elements to be found in the behavior set. Therefore, the task of designers is to create an institution that would effectively address such problem.

The two selections available under binary behavioral choice are: normal behavior b_1 and loafing b_2.

Table 7.5 Current institution parameters for an institution designed to tackle corruption at a coal inspection station (improved version)

Component name	Component type	Element	Numerical value or characteristic	Description
Behavior set	Binary behavior set	b_1	Normal coal management behavior	
		b_2	Truck diversions	
Promoter of b_1	Reward	r_1	RMB200,000/year * 5 years = RMB1,000,000	Assuming the station employs four workers, with annual wage per person at RMB50,000 and assuming a work period of five years
Observer of b_2	Equipment-type observer	p_{21}	0.99	Electronic monitoring and records, with probability of discovery at 99 %
		p_{22}	0.01	
Suppressor of b_2	Reward-type suppressor	s_2	RMB-200,000/year * 5 years + (RMB-1.2 million/year * 5 years) = RMB-7 million	Dismissal + jail term; assuming that the station employs four workers, with the loss to each staff member from dismissal standing at RMB-50,000/year and RMB-200,000/year for four members. Assuming that they could have worked another five years, total wage loss stands at RMB-1 million. The exchange utility for one person in jail is RMB-300,000/year and RMB-1.2 million/year for four persons. For a five-year jail term, total exchange utility stands at RMB-6 million
Promoter of b_2	Reward-type promoter	r_2	RMB200,000/year * 5 years + RMB100 million/year * 5 years = RMB501 million/year	Normal wages + bribes from letting "lighter trucks" pass, assuming such behavior can go on for five years
Utility of b_1		u_1	$u_1 = 100$	Assume that the individual under management is risk-neutral
Utility of b_2		u_2	$u_2 = 0.99 \times (-700) + 0.01 \times 50100 = -192$	Assume that the individual under management is risk-neutral

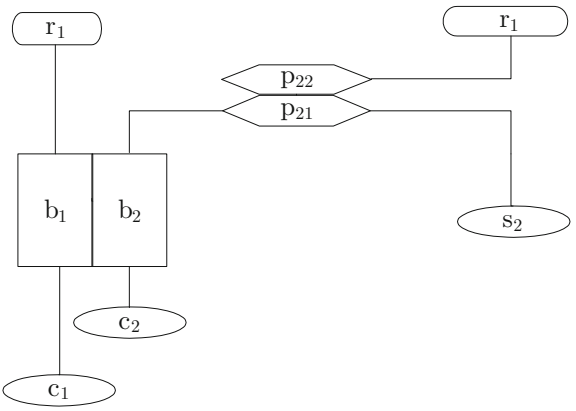

Fig. 7.6 Sun Diagram for institution designed to tackle problem of loafing

The penalty institution designed to address loafing is a special type of binary behavior management institution. When the phenomenon of loafing exists, institution parameters feature $(c_1 > c_2) \wedge (r_2 = r_1)$, and of which parameter $c_1 > c_2$ indicates that the cost of loafing c_2 is lower than the cost of normal behavior c_1. Further, parameter $r_2 = r_1$ indicates that if loafing has not been observed, the individual would be able to obtain reward r_1 commensurate with normal behavior, while "$\wedge$" indicates that the two conditions on the left and right sides of the symbol must exist simultaneously.

There are many instances of loafing: for example, there is are cases of workers putting in less effort than typically observed while earning a normal salary $r_2 = r_1$; there are also cases of students not studying hard but trying nevertheless to obtain results $r_2 = r_1$ (at this juncture, r_1 is the reward for normal behavior, such as "exams passed") not commensurate with the actual level of effort put in. r_2 is the reward to the individual when his or her loafing behavior has not been observed.

This characteristic tells us that if there is no behavior suppressor target at loafing, i.e., if there is no observer P_2 and penalty s_2 in place, the rational individual would definitely opt for loafing b_2 (Fig. 7.6).

7.4.3 Institution Parameters

Institution parameters are shown in Table 7.6.

Table 7.6 Institution parameters for penalty institution designed to tackle problem of loafing on the job

Behavior	Content	Reward	Probability	Cost	Utility
b_1	Normal behavior	r_1	1	c_1	$u_1 = r_1 - c_1$
b_2	Undesired behavior	r_1	p_{22}	c_2	$u_2 = p_{22}r_1 + p_{21}s_2 - c_2$
		s_2	p_{21}		$= (1 - p_{21})r_1 + p_{21}s_2 - c_2$

7.4.4 Institution Conditions

Institution conditions required include: the utility of normal behavior must be greater than the utility of loafing, i.e., $u_1 > u_2$. Therefore we have:

$$r_1 - c_1 > (1 - p_{21})r_1 + p_{21}s_2 - c_2$$

7.4.4.1 Requirements for Observation Intensity p_{21}

The above equation then gives us:

$$p_{21} > \frac{c_1 - c_2}{r_1 - s_2} \tag{7.10}$$

That is to say, if we are to consider institution observer performance, then the observational rate of the observer must meet requirements laid down in Eq. (7.10).

7.4.4.2 Requirements for the Difference $r_1 - s_2$ Between Income from Normal Behavior and Penalty for Undesirable Behavior

If we consider the difference $r_1 - s_2$ between the reward for normal behavior r_1 and the penalty for loafing s_2, then the difference must satisfy the requirements laid down in Eq. (7.11). (Note: the greater the value of s_2, the greater the reward from loafing; the smaller the value, the small the reward from loafing. We should also pay attention to penalties such as fines and banning of the caught candidates from future exams; the value of such rewards should be negative.)

$$r_1 - s_2 > \frac{c_1 - c_2}{p_{21}} \tag{7.11}$$

7.4.5 Case Study of a Penalty Institution Against Lazy Behavior: The Penalty Institution Against Cheating at Exams

7.4.5.1 Analysis of the Current Penalty Institution Against Cheating at Exams

(i) About the institution

The rules for punishing cheating students in exams by a district are as follows:

How minor misconduct at exams is to be handled: students caught copying exam answers and solutions would have the results of the exam in question voided.
How severe misconduct at exams is to be handled: candidates caught hiring proxies, cheating with communications devices during exams, etc., would have their results voided and banned from the exam for three years.

The institution implementers in this case are:

Human invigilators are in place to observe misconduct at the exams, with disciplinary measures meted out by the local Department of Education.

(ii) The Sun Diagram for the institution

The Sun Diagram for a penalty institution for severe misconduct during exams is shown as in Fig. 7.7.

(iii) Institution parameters

Parameters for a penalty institution for severe misconduct during exams are shown in Table 7.7.

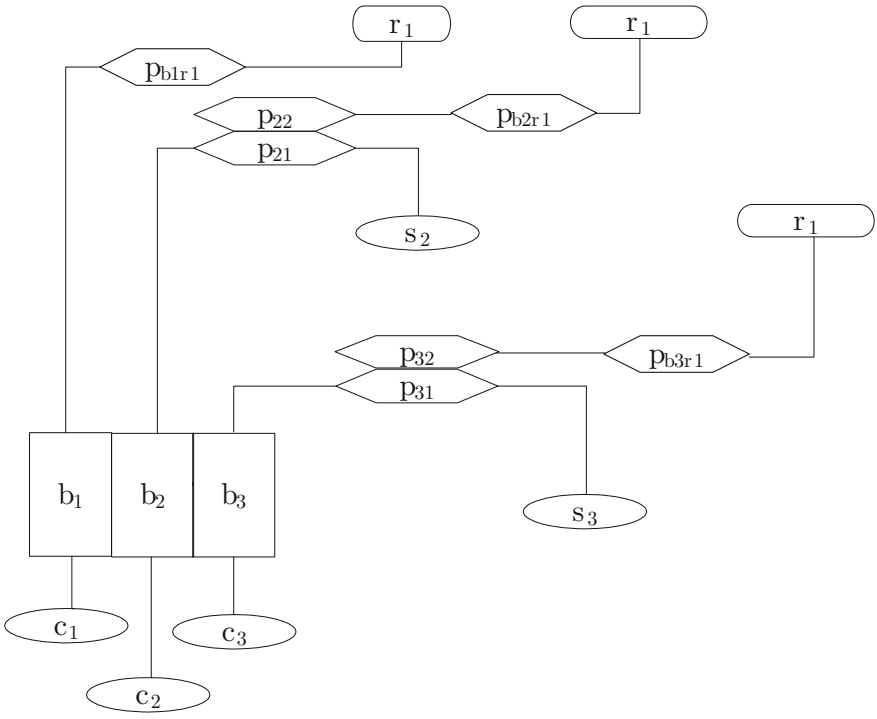

Fig. 7.7 The Sun Diagram for a penalty institution for severe misconduct during exams (ternary behavior)

(iv) Institution parameters

Of the parameter data listed above in the exam management institution, values r_1, s_2, s_3, c_1, c_2, c_3, etc., can come from statistical processing of student evaluations of the relevant indicators while p_{b1r1}, p_{21}, p_{31}, p_{b2r1} and p_{b3r1} can come from actual statistics. Here, the data used is fictional and only supplied for the purposes of illustrating the calculation procedure (Table 7.8).

(v) Calculating behavior priority for the individual

Utility calculation for each behavior (assuming individual is risk-neutral):

$$u_1 = p_{b1r1}r_1 - c_1 = 0.25 \times 25 - 4 = 2.25$$
$$u_2 = (1 - p_{21})p_{b2r1}r_1 + p_{21}s_2 - c_2 = (1 - 0.3) \times 0.3 \times 25 + 0.3 \times (-10) - 1 = 1.25$$
$$u_3 = (1 - p_{31})p_{b3r1}r_1 + p_{31}s_3 - c_3 = (1 - 0.3) \times 0.95 \times 25 + 0.3 \times (-25) - 2 = 7.125$$

The utility of each behavior is ranked in a descending order as follows:

$$u_3 > u_1 > u_2$$

Thus, under the current institution, the priority order for behavioral choices for the rational individual is:

$$b_3 \succ b_1 \succ b_2$$

That is to say, the institution is ineffective as the rational exam candidate (note that not all candidates are indeed rational in reality) would first opt for cheating behavior. We should note that this is only a theoretical result obtained from an institution mathematical model. In Chap. 8 of this book, we use the probability model to obtain results that are even closer to the situation in real life.

7.4.5.2 Institution Improvement I: Technical Measures

(i) Boundary conditions for effective institutions

First, we analyze the boundary conditions that render the institution effective. Based on the institution parameter table above, the equation for calculating behavioral utility for various behaviors is:

$$u_1 = p_{b1r1}r_1 - c_1$$
$$u_2 = (1 - p_{21})p_{b2r1}r_1 + p_{21}s_2 - c_2$$
$$u_3 = (1 - p_{31})p_{b3r1}r_1 + p_{31}s_3 - c_3$$

Table 7.7 Parameters for a penalty institution designed to tackle cheating during exams

Behavior	Content	Reward (two component levels)	Result and probability for level I	Result and probability for level II	Cost	Utility (Assuming individual is risk-neutral)
b_1	Normal behavior	r_1		Examination passed, p_{b1r1}	c_1	$u_1 = p_{b1r1}r_1 - c_1$
b_2	Minor misconduct	r_1	Minor misconduct not discovered, p_{22}	Examination passed, p_{b2r1}	c_2	$u_2 = p_{22}p_{b2r1}r_1 + p_{21}s_2 - c_2$ $= (1 - p_{21})p_{b2r1}r_1 + p_{21}s_2 - c_2$
		s_2	Minor misconduct discovered and punished, p_{21}			
b_3	Severe misconduct	r_1	Severe misconduct not discovered, p_{32}	Multiply the probability of receiving signals necessary for cheating by the probability of passing the exam, p_{b3r1}	c_3	$u_3 = p_{32}p_{b3r1}r_1 + p_{31}s_3 - c_3$ $= (1 - p_{31})p_{b3r1}r_1 + p_{31}s_3 - c_3$
		s_3	Severe misconduct discovered and punished, p_{31}			

Table 7.8 Institution parameters

Component	Component type	Numerical value
r_1	Examination passed	25
s_2	Exam results voided	−10
s_3	Exam results voided, barred from exams for three years	−25
c_1	Cost of studying hard: time and energy	4
c_2	Nervousness caused by misconduct	1
c_3	Nervousness caused by misconduct, and economic cost of paying for an exam proxy or purchase of communications equipment	2
p_{21}	Probability of discovering minor misconduct with human invigilators	0.3
p_{31}	Probability of discovering severe misconduct with human invigilators	0.3
p_{b1r1}	Probability of passing the exam the normal way	0.25
p_{b2r1}	Probability of passing the exam with minor misconduct not discovered	0.3
p_{b3r1}	Probability of passing the exam with severe misconduct not discovered	0.95

This way, if we want the rational candidate to prioritize normal exam behavior b_1, then the following conditions must be satisfied at the same time:

First, $u_1 > u_2$:

i.e., $p_{b1r1}r_1 - c_1 > (1 - p_{21})p_{b2r1}r_1 + p_{21}s_2 - c_2$

We get:

$$p_{21} > \frac{(p_{b2r1} - p_{b1r1})r_1 + (c_1 - c_2)}{(p_{b2r1}r_1 - s_2)}$$

Substitute the values from the institution parameter table above and determine the smallest probability boundary value p_{31} required for the institution to be effective:

$$p_{21} > \frac{(p_{b2r1} - p_{b1r1})r_1 + (c_1 - c_2)}{(p_{b2r1}r_1 - s_2)} = \frac{(0.3 - 0.25)25 + (4 - 1)}{(0.3 \times 25 - (-10))} = 0.24$$

Then, $u_1 > u_3$:

i.e., $p_{b1r1}r_1 - c_1 > (1 - p_{31})p_{b3r1}r_1 + p_{31}s_3 - c_3$.

We get:

$$p_{31} > \frac{(p_{b3r1} - p_{b1r1})r_1 + (c_1 - c_3)}{(p_{b3r1}r_1 - s_3)}$$

Substitute the values from the institution parameter table above and determine the smallest probability boundary value p_{31} required for the institution to be effective:

$$p_{31} > \frac{(p_{b3r1} - p_{b1r1})r_1 + (c_1 - c_3)}{(p_{b3r1}r_1 - s_3)} = \frac{(0.95 - 0.25)25 + (4 - 2)}{(0.95 \times 25 - (-25))} = 0.4$$

This way, if other institution parameters apart from that of the observer do not change, the condition for institution effectiveness would be: the parameters for both p_2 and p_3 must satisfy the following at the same time:

$$(p_{31} > 0.4) \wedge (p_{21} > 0.24)$$

(ii) Improvement measures

Comparing the conditions above with the component parameters from the original institution, we see that condition $p_{21} > 0.24$ has been met by the original institution as $p_{21} = 0.3$. The main problem lies with $p_{31} = 0.3 < 0.4$. Hence, we would need to raise the value of p_{31}. The following measures were implemented as a result.

Measure I (improving observer performance)

As for candidates cheating at exams with the help of communications devices, it was decided that observer performance be boosted with the installation of equipment that detects electronic signals. This way, observer performance or p_{31} would be enhanced significantly.

At the same time, to lower the utility of cheating behavior b_3, the following measure can also be taken.

Measure II (lowering the expected reward from cheating behavior)

Electronic signal jammers were used to significantly lower the expected reward from using communications devices to cheat. In other words, p_{b3r1} was diminished significantly.

The aforementioned two measures show that in the course of institution design that the use of high-performance technical equipment can significantly improve institution effect.

(iii) Sun Diagram following institution improvement

As the institution structure in the Sun Diagram remains unchanged, the Sun Diagram for the institution post-improvement is the same as is shown in Fig. 7.7.

(iv) Institution parameters following institution improvement

As the institution structure in the Sun Diagram remains unchanged, institution parameters have also remained unchanged. Therefore, the institution parameter table is as shown in Table 7.7.

Table 7.9 Parameters for a penalty institution designed to tackle cheating during exams (post-institution improvement)

Component	Component type	Numerical value
r_1	Result obtained	25
s_2	Exam results voided	−5
s_3	Exam results voided, barred from exams for three years	−25
c_1	Cost of studying hard: time and energy	4
c_2	Nervousness caused by misconduct	1
c_3	Nervousness caused by misconduct, and economic cost of paying for an exam proxy or purchase of communications equipment	2
p_{21}	Probability of discovering minor misconduct with human invigilators	0.3
p_{31}	Probability of discovering severe misconduct using human invigilators + electronic signal detectors	0.9
p_{b1r1}	Probability of passing the exam the normal way	0.25
p_{b2r1}	Probability of passing the exam with minor misconduct not discovered	0.3
p_{b3r1}	Multiply the probability of receiving electronic signals for cheating 0.1 by the probability of not being caught and passing the examples 0.95	0.095

(v) Component data following institution improvement

However, as certain technical measures were deployed, there has been a change in institution component performance. Therefore, the data for institution components has also changed.

The parameters for a penalty institution designed to tackle cheating during exams (post-institution improvement) is as shown in Table 7.9.

(vi) Calculating behavior priority for individuals

Utility calculation for each behavior (assuming individual is risk-neutral):

$$u_1 = p_{b1r1}r_1 - c_1 = 0.25 \times 25 - 4 = 2.25$$
$$u_2 = (1 - p_{21})p_{b2r1}r_1 + p_{21}s_2 - c_2 = (1 - 0.3) \times 0.3 \times 25 + 0.3 \times (-10) - 1 = 1.25$$
$$u_3 = (1 - p_{31})p_{b3r1}r_1 + p_{31}s_3 - c_3 = (1 - 0.9) \times 0.095 \times 25 + 0.9 \times (-25) - 2 = -26.26$$

The utility of each behavior is ranked in descending order as follows:

$$u_1 > u_2 > u_3$$

Following institution improvement, rational candidates (note that not all candidates are rational in reality) would prioritize behavior options as follows:

$$b_1 \succ b_2 \succ b_3$$

7.4.5.3 Institution Improvement II: Management Measures

(i) Methods

An analysis of the penalty institution against cheating in exams shows that the original institution was comparatively ineffective. A key reason for this situation is that the probability of success in cheating by using communications devices was rather high. This was why quite a few students had opted for this method.

As such, managers can try to control behavioral resources for cheating behavior b_3, such as prohibiting the sales of such resources in order to lower the probability of students obtaining such devices. This in turn would lower the probability of cheating behavior occurring, with $p(b_3)$ diminished significantly.

This is a simple management measure. What this illustrates is that pure management improvements (as opposed to technical improvements) can also lead to improved institution performance.

(ii) Sun Diagram for penalty institution designed to tackle cheating during exams (post-institution improvement) (Fig. 7.8)
(iii) Parameters for a penalty institution designed to tackle cheating during exams (post-institution improvement) (Table 7.10)
(iv) Parameters for a penalty institution designed to tackle cheating during exams (post-institution improvement) (Table 7.11)
(v) Calculating behavior priority for the individual.

As the improvement measures taken work mainly through the limiting of behavioral resources in order to change the probability $p(b_3)$ of cheating behavior b_3 occurring, the various parameters used in the calculation of behavioral utility remain unchanged. Thus, after post-institution improvement, the rational candidate's order of priority for various behavioral options remains the same, i.e., $b_3 \succ b_1 \succ b_2$.

(vi) Probability of occurrence for each behavior

However, in contrast to the original institution, due to the limit on behavioral resources the probability of cheating behavior b_3 occurring is a mere 0.3 compared to the previous 1. Hence, the probability of occurrence for various behaviors in the behavior set would be (when the individual is rational):

The probability of severe misconduct during exams b_3 occurring is $p(b_3) = 0.3$.
The probability of minor misconduct during exams b_2 occurring is $p(b_2) = 0$.

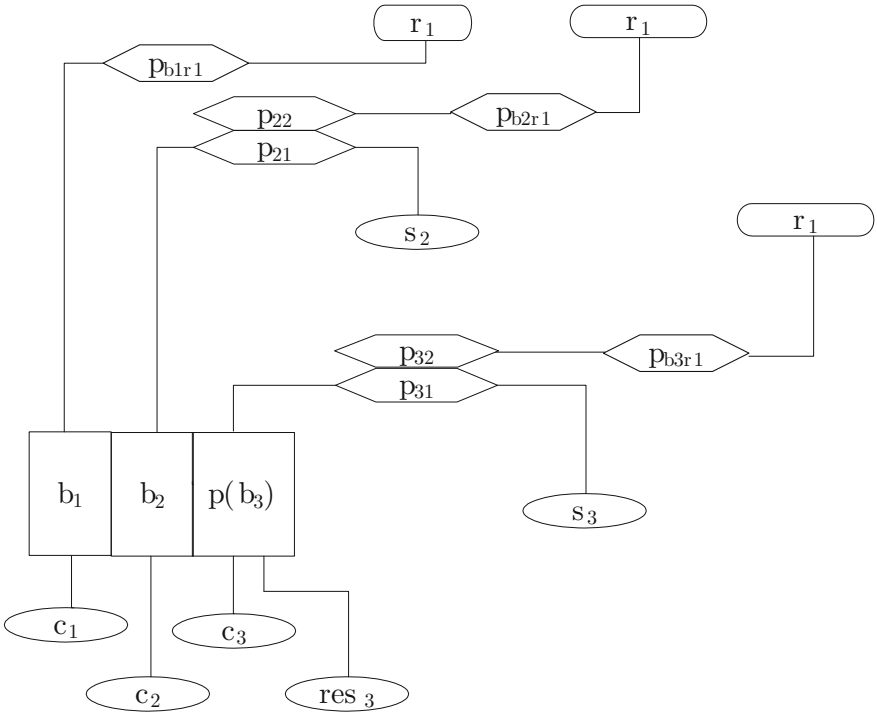

Fig. 7.8 Sun Diagram for a penalty institution designed to tackle cheating during exams (post-institution improvement)

The probability of normal behavior during exams b_1 occurring is $p(b_1) = 1 - p(b_3) = 1 - 0.3 = 0.7$.

I.e., the probability vector for the behavior set would be:

$$p(B) = (0.7, 0, 0.3)$$

This is because the rational exam candidate (note that not all candidates are rational in reality) would opt for cheating if possible. However, as the measure to ban the sales of enabling equipment has been implemented, the probability of obtaining such equipment for cheating has decreased significantly to 0.3. Hence, the probability of occurrence for cheating behavior b_3 is 0.3.

If cheating behavior b_3 cannot be realized, then candidates would then go down the priority list to the next option, normal exam behavior b_1. This means that the probability of such behavior occurring, b_1, is 0.7.

As the utility from b_2 is lowest, and at the same time as utility for behavior b_1 (which is not restricted) is higher than b_2, the rational candidate (note that not all candidates are rational in reality) would definitely not opt for b_2, making the probability of such behavior occurring 0.

Table 7.10 Parameters for a penalty institution designed to tackle cheating during exams (post-institution improvement)

Behavior	Content	Probability	Reward (two component levels)	Result and probability for level I	Result and probability for level II	Cost	Utility (assuming individual is risk-neutral)
b_1	Normal behavior	1	r_1		Examination passed, p_{b1r1}	c_1	$u_1 = p_{b1r1}r_1 - c_1$
b_2	Minor misconduct	1	r_1	Minor misconduct not discovered, p_{22}	Examination passed, p_{b2r1}	c_2	$u_2 = p_{22}p_{b2r1}r_1 + p_{21}s_2 - c_2$ $= (1 - p_{21})p_{b2r1}r_1 + p_{21}s_2 - c_2$
			s_2	Minor misconduct discovered and punished, p_{21}			
b_3	Severe misconduct	$p(b_3)$	r_1	Severe misconduct not discovered, p_{32}	Multiply the probability of receiving signals necessary for cheating by the probability of passing the exam, p_{b3r1}	c_3	$u_3 = p_{32}p_{b3r1}r_1 + p_{31}s_3 - c_3$ $= (1 - p_{31})p_{b3r1}r_1 + p_{31}s_3 - c_3$
			s_3	Severe misconduct discovered and punished, p_{31}			

Table 7.11 Parameters for a penalty institution designed to tackle cheating during exams (post-institution improvement)

Component	Component type	Numerical value
r_1	Result obtained	25
s_2	Exam results voided	−10
s_3	Exam results voided, barred from exams for three years	−25
c_1	Cost of studying hard: time and energy	4
c_2	Nervousness caused by misconduct	1
c_3	Nervousness caused by misconduct, and economic cost of paying for an exam proxy or purchase of communications equipment	2
p_{21}	Probability of discovering minor misconduct with human invigilators	0.3
p_{31}	Probability of discovering severe misconduct with human invigilators	0.3
p_{b1r1}	Probability of passing the exam the normal way	0.25
p_{b2r1}	Probability of passing the exam with minor misconduct not discovered	0.3
p_{b3r1}	Probability of passing the exam with severe misconduct not discovered	0.95
$p(b_3)$	Probability of b_3 occurring with the ban on the sales of communications equipment used for cheating	0.3

7.5 Bonus Institution of Binary Behavioral Choices

7.5.1 Basic Features of the Bonus Institution of Binary Behavioral Choices

A bonus institution seeks to encourage actors to opt for desirable behavior. The bonus institution is relatively common in human society and has a wide range of application.

The distinctive feature of a binary or two-dimensional bonus institution is that there is only one individual under management and this individual has to choose between one of two possible behaviors within a binary behavior set: b_1 or normal behavior and b_2, desirable behavior.

The institution does not monitor for the occurrence of the individual's normal behavior b_1. Rather, as long as the individual opts for normal behavior b_1, he will receive reward r_1. Hence, we see that in this institution there is necessarily a relationship between normal behavior b_1 and the reward r_1.

A two-dimensional observer p_2 is deployed for desirable behavior b_2. When desirable behavior b_2 occurs, the observer p_2 would be able to detect b_2 at a probability of p_{21}. At this juncture, the individual will receive reward r_2 for choosing b_2. Clearly, within the bonus institution, $r_2 > r_1$. Meanwhile, as the

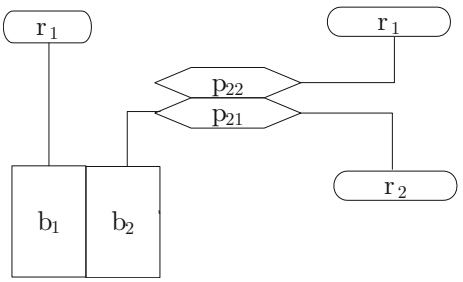

Fig. 7.9 Sun Diagram for bonus institution of binary behavioral choice

observer p_2 may not work 100 % of the time, when desirable behavior b_2 occurs, the observer p_2 will malfunction, i.e., fail to observe b_2 at a probability of $p_{22} = 1 - p_{21}$. At this point, the individual can only obtain the reward for normal behavior r_1.

7.5.2 Sun Diagram for Bonus Institution of Binary Behavioral Choice

The Sun Diagram for a bonus institution of binary behavioral choice is shown as in Fig. 7.9.

The fundamental feature of the Sun Diagram for the bonus institution of binary behavioral choice is that there is a binary behavior set that comprises normal behavior b_1 and desirable behavior b_2. Of the two, normal behavior b_1 is directly linked to the promoter r_1. Desirable behavior b_2 is directly linked to the binary observer p_2, with the "observation output (observed)" of the observer linked to promoter r_2 while "observation output (not observed)" is linked to the promoter r_1.

In terms of the Sun Diagram structure, the reward and penalty institutions are structurally similar. The difference between the two is that the observer in the penalty institution is targeted at undesirable behavior while the observer in the bonus institution is targeted at desirable behavior.

7.5.3 Institution Parameters and Conditions

Parameters for a bonus institution of binary behavioral choice are shown in Table 7.12.

The condition for a binary or two-dimensional bonus institution is $u_1 < u_2$, i.e., $r_1 - c_1 < (1 - p_{21})r_1 + p_{21}r_2 - c_2$.

This gives us:

$$p_{21} > \frac{c_2 - c_1}{r_2 - r_1} \tag{7.12}$$

Table 7.12 Parameters for a bonus institution of binary behavioral choice

Behavior	Content	Reward	Probability	Cost	Utility
b_1	Normal behavior	r_1	1	c_1	$u_1 = r_1 - c_1$
b_2	Desirable behavior	r_1	p_{22}	c_2	$u_2 = p_{22}r_1 + p_{21}r_2 - c_2$
		r_2	p_{21}		$= (1 - p_{21})r_1 + p_{21}r_2 - c_2$

That is to say, if observation intensity p_{21} is lesser than the requirements set out in Eq. (7.12), the institution is ineffective and the rational individual would opt for normal behavior b_1 rather than desirable behavior b_2.

When transforming the above equation, we get:

$$r_2 > \frac{c_2 - c_1 + p_{21}r_1}{p_{21}} \tag{7.13}$$

This means that with a specific observation intensity level p_{21}, rewards for desirable behavior must meet the requirements set down in Eq. (7.13) for the institution to be effective. That is to say, it is only at this point that the rational individual would opt for desirable behavior.

7.5.4 Applicability of Penalty and Bonus Institutions

The penalty institution is a management institution that is commonly used. For instance, the case of appointees considered to have completed their terms successfully if no major incidents have occurred during his tenure is an example of a penalty institution. This type of institution is especially useful for desirable behavior that is more challenging to observe compared to undesirable behavior. For example, the work performance of a public servant can be hard to monitor and observe. Hence, his wage levels cannot be determined by workload. The only thing his superiors can do is to observe for obvious instances of dereliction of duty. When such instances have not been observed (even if they have occurred), the public servant would be treated as having acted normally and will receive his wages as per normal.

The bonus institution is a management institution that is also commonly used. For example, research institutions typically implement bonus institutions with their researchers while enterprises also implement bonus institutions (e.g., commission institutions) with their sales personnel. This type of institution is especially useful in situations where undesirable behavior is more challenging to observe compared to desirable behavior. For example, researchers "loafing on the job" can be hard to

detect, and this is the same with salespersons who do not put in their all making sales calls. Under such circumstances, the institution manager can only observe for clearly desirable behavior (or their proxies), such as research output or sales performance, and then grant rewards based on such results.

Reference

Zhang C (2010) Unveiling the Qing Imperial Household Department: the most corruption-ridden agency in ancient China. Ifeng Books. Retrieved from: http://book.ifeng.com/shuzhai/detail_2010_05/15/1520542_2.shtml. 15 May 2010

Chapter 8
Behavior Probability Under the Behavioral Management Institution

8.1 The Definition and Significance of Behavior Probability

8.1.1 Definition of Behavior Probability

There are two definitions for the term "behavior probability", the first being: the chance of a certain behavior occurring in a group that is being managed. For instance, if 10 individuals in a group of 100 persons opt for a certain behavior b_i, then the probability of that behavior would be $p(b_i) = 0.1$.

The second scenario is centered on the individual who has to make a certain behavioral choice on multiple occasions. Here, the term Behavior probability refers to the ratio of a certain behavioral choice out of the total number of times the individual has had to make that choice. For example, a student has to make a choice about what to do with his time after dinner. Over the course of one month, he chose to revise his homework on 20 occasions and to head to the Internet cafe on another 10 occasions. Therefore, the probability for "homework revision" would be

$$p(\text{homework revision}) = \frac{20}{20+10} = 0.67$$

while the probability for "heading to the Internet cafe" would be

$$p(\text{heading to the Internet cafe}) = \frac{10}{20+10} = 0.33.$$

It should be noted that in actual institution design, as the objectives of the behavior management mechanism may vary and as the results of such a mechanism are often more observable than the choice itself, the calculation of behavior probability is often performed using the relevant behavioral results as variables. For example, the proportion of savings deposits to total disposable income can be used

S. Sun, *Five Basic Institution Structures and Institutional Economics*,
DOI 10.1007/978-981-10-0344-8_8

in an institution designed to manage economic behavior to indicate the probability of "savings behavior"). In a tourism-related management institution, the proportion of tourists traveling to Spot A to the total number of travelers can be used to indicate the behavior probability for "traveling to Spot A", etc.

8.1.2 Significance of Behavior Probability

8.1.2.1 The Group and Repeatable Nature of Behavior

Behavior probability has great significance in the design of an institution. This is because in general, an institution is designed for a group of people rather than for an individual. In other words, an institution is designed to manage a target group. Similarly, an institution is not designed for one-off occurrences but for behavior that happens again and again over time. These two types of situations allow behavior probability to be used as a descriptor for behavior choice.

8.1.2.2 Managing Differences Between Individuals in the Target Group

The management of social behavior tells us that the behavior of individuals will not be uniform. Some of the behaviors exhibited are high-utility behaviors with slightly higher probability of occurring, while some low-utility behaviors—even those with negative utility—persist to some degree.

Cognitive gaps constitute one of the key reasons for the difference in behavioral choice from person to person. Factors of irrationality can influence individuals' cognition of behavioral utility to some extent, which means that different individuals can perceive the same thing quite differently.

Behavioral gaps are the second key reason for the difference in behavioral choice from person to person. Even if everyone in the group perceives behavioral utility in the same way, as individual control over behavior varies from person to person, the behavioral choices that occur will also differ. For example, the harms of drug-taking are well known: loss of personal property and penalties from the state laws. This is why most members of society would not engage in the act of drug consumption. However, no matter how severe the penalties are in a given society, there will always be a proportion of individuals who do engage in the act.

The third key factor that impacts behavioral choice is the fact that the same behavior may have a different utility for different persons. The specific environments that individuals find themselves in vary, that's why similar behavioral outcomes may pose vastly different utilities from person to person. For example, for a starving man the most "high-utility" behavior would be to look for food, while for a rich man the most "high-utility" behavior would be to play a game of golf.

The fourth reason for behavioral gaps is the cost difference caused by behavioral paths. Past behavioral choices by individuals form what is a "behavioral path". If two persons are to make the same behavioral choice, their respective behavioral paths would make the cost of that choice different. This cost gap in turns causes behavioral utility to vary from individual to individual. For instance, it is much easier for a biology major in college to switch to the study of medicine compared to a history major wishing to do the same. In real-life conditions, the knowledge and experiences, wealth, social standing, access to social networks, etc., of individuals all form advantages for specific behaviors. If the individual were to choose not to make use of these advantages and select an inappropriate behavior, he will incur a much greater cost for himself and in turn generate reduced behavioral utility.

Therefore, although in some studies individuals' judgments of benefits have been regarded as "uniform", it is worth remembering that this is but an approximation of actual circumstances in order to simplify the study process. In fact, there are a multitude of factors that have an impact on individuals' behavioral choices, and for different individuals these factors can also vary greatly, thereby leading to a large number of behavioral choices.

With such a situation, the probability of the occurrence of a given behavior within a group serves as a more holistic reflection of the differences between the individuals within as well as of the overall impact of various factors, both rational and irrational. This makes behavior probability an accurate indicator to describe behavioral choices made by individuals.

8.1.2.3 Behavior Probability Is a Key Indicator for Measuring Actual Effect of Institution Design

Whether in organizational or social management, an institution is often required to regulate behaviors of individuals. The chief indicator of the effect of a given institution design is whether the behavior probability measured matches the expected probability. This is because in the case where the individual targets of management belong to a social grouping, no matter how effective the institution is it is impossible to achieve the scenario wherein "everyone opts for the desirable behavior". Therefore, so long as the behavior probability for the desirable behavior has been raised we can regard the institution to have positive effect. Similarly, no matter how ineffective the institution is, the situation wherein "everyone opts for the undesirable behavior" will not occur. As long as the behavior probability of the undesirable behavior has not been lowered, the institution is ineffective. Therefore, behavior probability is a key indicator for measuring actual institution effect.

8.2 Behavior Probability Model Based on Behavioral Utility

8.2.1 The Basic Model

The behavior set of all actors contains n behaviors, with the behavior set satisfying both uniqueness and completeness criteria for behavior choice. This means that the actor must select one of the behaviors within b_i with the behavior probability model expressed in Eq. (8.1).

$$\left.\begin{aligned} p(b_1) &= \frac{e^{\alpha_1+\beta_1 u_1}}{e^{\alpha_1+\beta_1 u_1}+\cdots+e^{\alpha_i+\beta_i u_i}+\cdots+e^{\alpha_n+\beta_n u_n}} \\ &\vdots \\ p(b_i) &= \frac{e^{\alpha_i+\beta_i u_i}}{e^{\alpha_1+\beta_1 u_1}+\cdots+e^{\alpha_i+\beta_i u_i}+\cdots+e^{\alpha_n+\beta_n u_n}} \\ &\vdots \\ p(b_n) &= \frac{e^{\alpha_n+\beta_n u_n}}{e^{\alpha_1+\beta_1 u_1}+\cdots+e^{\alpha_i+\beta_i u_i}+\cdots+e^{\alpha_n+\beta_n u_n}} \\ \text{或者} p(b_n) &= 1-p(b_1)-\cdots-p(b_i)-\cdots-p(b_{n-1}) \end{aligned}\right\} \tag{8.1}$$

In the model, $p(b_i)$ is the behavior probability for b_i, while u_i is the expected utility of behavior b_i. α_i and β_i are regression coefficients obtained based on historical data with a value interval of $\alpha_i \in (-\infty, \infty)$, $\beta_i > 0$, $i \in [1, 2, \ldots, n]$.

From the model, by assuming the expected utility of other behaviors remains unchanged:

$$\left.\begin{aligned} \lim_{u_i \to \infty} p(b_i) &= 1 \\ \lim_{u_i \to -\infty} p(b_i) &= 0 \end{aligned}\right\}$$

we can see that the range for $p(b_i)$ does meet probability requirements.

8.2.2 Significance of Parameters

For $p(b_i) = \dfrac{e^{\alpha_i+\beta_i u_i}}{e^{\alpha_1+\beta_1 u_1}+\cdots+e^{\alpha_i+\beta_i u_i}+\cdots+e^{\alpha_n+\beta_n u_n}}$, the parameters α_i and β_i do carry a degree of significance.

First, consider $\alpha_i \in (-\infty, \infty)$. It determines when the expected utility of all behaviors in the set is zero, i.e., $u_1 = u_2 = \cdots = u_i = \cdots = u_n = 0$, the probability of occurrence for b_i is recorded as $p_0(b_i)$:

$$p_0(b_i) = \frac{e^{\alpha_i}}{e^{\alpha_1} + \cdots + e^{\alpha_i} + \cdots + e^{\alpha_n}}$$

This probability occurs without any motivation or incentive, and is called the "baseline" probability b_i. Typically, the factors that have the greatest impact on baseline probability are the behavioral habits of individuals and the state of their behavioral paths.

The factors that impact β_i are dominant knowledge and beliefs in the group, cognitive capability of the majority of individuals within, and their cognitive gap. Therefore β_i is an overall reflection of the factors that affect the behavior probability.

"Knowledge and beliefs" include the knowledge required to determine utility and the structure of personal values (such as the individual actor's proclivity towards factors like economic benefit, reputation and social standing and appetite for risk). "Cognitive capability" (including the ability for judgment) and cognitive gaps also have an impact on β_i.

When using the model, we should first make use of the relevant regression coefficients derived from the historical data. With behavior probability, we can make use of the corresponding statistics while for utility, in the case of the economic behavior we may make use of the corresponding historical data (by assuming that the individual is risk-neutral, we may directly use expected benefits to stand for utility). For the non-economic behavior, we can make use of exchange utility and with economic benefit as the basis for investigating the estimated value put to the exchange utility of the behavior in question by actors at that time.

After obtaining the regression coefficients for the model, we may then use the model to forecast the impact of the model on institution design utility. In other words, we bring in the utility of various behaviors under the institution into the model to estimate the behavior probability of each respective behavior.

8.3 Example of Calculating Behavior Probability

Now, with the utility-based behavior probability model, we estimate the behavior probability of cheating during exams in the example used by Sect. 7.4 of Chap. 7 for both "before institution refinement" and "after institution refinement" scenarios.

8.3.1 Probability of Cheating in Exams Before the Improving of the Penalty Institution

8.3.1.1 Behavior Set of the Managed

Individuals being managed by the institution can choose from the three behavioral options:

b_1 Take the exams in a normal, compliant fashion;
b_2 Take the exams in a non-compliant fashion, such as answering with copied solutions and answers (milder offenses);
b_3 Ask someone else to take the exams, or use telecommunications equipment during the exams, and commit other severe offenses.

8.3.1.2 Analysis of the Pre-improved Penalty Institution for Cheating in Exam Under Hypothesis of Rational Man

Based on the calculations in Chap. 7, the utility of the pre-improvement institution for the respective behaviors are: $u_1 = 2.25$, $u_2 = 1.25$ and $u_3 = 7.125$.

The utility of each behavior is ranked in a descending order as follows: $u_3 > u_1 > u_2$.

Thus, under the pre-improvement institution, the priority order for behavioral choices that could be made by individuals is: $b_3 \succ b_1 \succ b_2$. In other words, the results of analysis based on the theory of the institution model laid out in Chap. 7 is that if all candidates are rational they would choose to cheat on their exams.

8.3.1.3 Using the Behavior Probability Model to Estimate Behavior Probability

A Behavior Probability Model that Tackles Cheating Behaviors During Exams

With regard to the three types of behaviors possible, the behavior probability model is as follows:

$$p(b_1) = \frac{e^{\alpha_1 + \beta_1 u_1}}{e^{\alpha_1 + \beta_1 u_1} + e^{\alpha_2 + \beta_2 u_2} + e^{\alpha_3 + \beta_3 u_3}}$$
$$p(b_2) = \frac{e^{\alpha_2 + \beta_2 u_2}}{e^{\alpha_1 + \beta_1 u_1} + e^{\alpha_2 + \beta_2 u_2} + e^{\alpha_3 + \beta_3 u_3}}$$
$$p(b_3) = \frac{e^{\alpha_3 + \beta_3 u_3}}{e^{\alpha_1 + \beta_1 u_1} + e^{\alpha_2 + \beta_2 u_2} + e^{\alpha_3 + \beta_3 u_3}}$$

Coefficients in the Behavior Probability Model

Assume that the coefficients for each are as follows after regression of historical data:

$$\alpha_1 = -2.5903, \quad \beta_1 = 0.9733$$
$$\alpha_2 = -3.7297, \quad \beta_2 = 1.6119$$
$$\alpha_3 = -6.9078, \quad \beta_3 = 0.7033$$

Therefore, the numerical value of the model is as follows:

$$p(b_1) = \frac{e^{-2.5903+0.9733u_1}}{e^{-2.5903+0.9733u_1} + e^{-3.7297+1.6119u_2} + e^{-6.9078+0.7033u_3}}$$
$$p(b_2) = \frac{e^{-3.7297+1.6119u_2}}{e^{-2.5903+0.9733u_1} + e^{-3.7297+1.6119u_2} + e^{-6.9078+0.7033u_3}}$$
$$p(b_3) = \frac{e^{-6.9078+0.7033u_3}}{e^{-2.5903+0.9733u_1} + e^{-3.7297+1.6119u_2} + e^{-6.9078+0.7033u_3}}$$

Calculating the Current Behavior Probability

Substitute data $u_1 = 2.25$, $u_2 = 1.25$ and $u_3 = 7.125$ in the above equation to calculate behavior probability for b_1, b_2 and b_3, there will be

$$p(b_1) = 0.67, \quad p(b_2) = 0.18 \quad \text{and} \quad p(b_3) = 0.15.$$

Hence, we see that despite rational analysis the expected utility of $u_3 > u_1 > u_2$ or cheating at the exams is still the greatest. As such, the priority order for possible behaviors should be $b_3 \succ b_1 \succ b_2$. However, due to the effect of beliefs, attitudes and others, the majority of actors chose to take the exam in the normal fashion (67 % of all candidates) with 18 % of candidates committing minor offenses. Although the reward with committing severe offenses is the greatest, only 15 % of all candidates chose to go down this path (although this is quite a high ratio still).

8.3.2 Probability for Cheating in Exams After Institution Refinement

8.3.2.1 Various Behavioral Utilities Within an Improved Institution

Two measures were implemented to refine the institution: first, the detection of electronic signals was stepped up as a means of enhancing exam invigilation and observer performance. Second, jamming signals were deployed to reduce the expected reward of severe offenses on the part of actors, thereby significantly reducing the expected reward of the use of telecommunications devices in the exam hall (for more details, please refer to Chap. 7).

Following the implementation of these measures, the respective utility of each behavior becomes: $u_1 = 2.25$, $u_2 = 1.25$, and $u_3 = -26.2625$. Thus, the utility of each behavior is ranked in a descending order as follows: $u_1 > u_2 > u_3$.

This way, following the refinement of the institution, if individuals are fully rational they will prioritize the various behaviors as $b_1 \succ b_2 \succ b_3$, i.e., all of them would opt for taking the exam in the normal fashion with zero occurrence of offenses.

8.3.2.2 Actual Behavior Probabilities for Various Behaviors Within an Improved Institution

Substitute the various utilities derived from the calculations after the institution improvement ($u_1 = 2.25$, $u_2 = 1.25$ and $u_3 = -26.2625$) into the behavior probability equation to obtain the following probabilities:

$$\begin{aligned} p(b_1) &= \frac{e^{-2.5903+0.9733u_1}}{e^{-2.5903+0.9733u_1} + e^{-3.7297+1.6119u_2} + e^{-6.9078+0.7033u_3}} \\ &= \frac{e^{-2.5903+0.9733\times 2.25}}{e^{-2.5903+0.9733\times 2.25} + e^{-3.7297+1.6119\times 1.25} + e^{-6.9078+0.7033\times(-26.2625)}} \\ &= 0.79 \end{aligned}$$

$$p(b_2) = 0.21$$

$$p(b_3) = 0.00000000001$$

It is clear that under the refined exam management institution, the percentage of individuals who chose to take the exams in a compliant manner rose from 67 to 79 % while the ratio of those who chose to commit severe offenses has declined to one in a trillion, or essentially 0. However, the percentage of individuals who opted for mild offenses rose from 18 % under the original institution to 21 % under the new institution. This can be explained thusly: as the committing of serious offense no longer provides the possibility of success under the new institution, a number of candidates otherwise inclined to committing severe offenses turned to taking the exams normally while the other part of this group (particularly those who would not be able to pass their exams if they do not cheat) turned to committing milder offenses such as peeking at another candidate's exam papers.

Here, we see irrational decision-making at work again. As $u_1 = 2.25$ and $u_2 = 1.25$, $u_1 > u_2$, if a candidate is completely rational, he or she would select b_1 all of the time. However, in reality, a good number of candidates selected b_2. The preceding shows the practical value of a behavior probability model in the estimation of institution design effect.

Chapter 9
Task Sharing Institutions

This chapter and Chap. 10 are different from Chap. 7 in that we will look specifically at the impact of institution structure on institution effect rather than the impact of institution component performance on institution effect. Hence, apart from using the Sun Diagram to illustrate the structure of various institution types, we will also analyze individual effort equilibrium levels with the help of mathematical models. The key objective of this type of analysis is to compare individual effort levels in various institutions with precision so that we can identify the factors that affect effort levels, and understand their respective effects. At the same time, we will also analyze certain institutions tweaked to regulate individual effort levels in order to yield certain principles of this process. These analyses will allow us to better understand the impact of institution structure on individual effort levels and provide us with specific tips on how to improve various institution designs.

9.1 The Basis for Comparing Effort Levels of Respective Individuals Under Various Task Sharing Institutions: Group-Optimal Levels

The task sharing institution is an institution where multiple individuals in a group work to complete a certain task. Within a manufacturing enterprise or public institution, the task sharing institution is require to ensure that members work together to complete the stipulated task.

The core questions to be studied with the task sharing institution are, the equantequation of effort level of individual, and the ways to optimize the equantequation.

In the task sharing institution, if the behavior in question is production behavior and the individual involved is an enterprise, then behavioral effort level can be indicated by the firm's production scale.

As individual task sharing institutions can vary in terms of their specific features, the divisibility of behavioral output would also vary. Further, as the divisibility of

S. Sun, *Five Basic Institution Structures and Institutional Economics*,
DOI 10.1007/978-981-10-0344-8_9

behavioral cost can also differ from specific institution to institution, we can categorize the task sharing institution into three sub-types: the double-independent institution, the reward-sharing institution, and the cost-pooling institution.

The basis for comparing the effect of these three types of the task sharing institution is the scale of total group benefits when individual effort level reaches equilibrium. No matter which type of institution involved, the institution that produces the highest overall group benefit is the best institution.

As the task sharing institution assigns specific sections of the job to various group participants, these participants must work together to complete the task. For the convenience of analysis, we assume the following two conditions for the group:

First, we assume that the reward of the group is independent from its environment. This means that the sum total of all behavioral rewards received by group participants is equal to total group output, and that such rewards originate solely from group output with no remainder left or "loss" to the group. That is, when total rewards are distributed within the group, no "further support" is required from other sources, nor is any "tribute" or payment to any party outside the group required.

Second, we assume that individuals in the group are of the same nature with no differences from individual to individual. This means that the analytic results for an individual within a specific institution will applicable to all other participants within the institution. This assumption frees us from the bother of analyzing individuals within the institution one by one. So long as we study one individual within the institution when conducting institution analysis, we would be able to understand all individuals within the institution.

These two assumptions help to make the analysis of institution structure simpler and more straightforward.

9.1.1 Individual's Output Function

Group output is the sum total of the individual output of all individuals. Therefore, we would need to first create the output function for the individual.

Assume that there are n individuals in the group, with the effort level of individual i being e_i. The relationship function between individual i's output r_i and effort level e_i is as in Eq. (9.1):

$$r_i = r_0 - \frac{r_0}{1 + e_i} \tag{9.1}$$

In the equation, $r_0 \geq 0, e_i \geq 0$, with r_0 the maximum individual output, such as the maximum production capacity for a factory, the highest catch possible for a fishing boat, the maximum agricultural output possible for a specific plot of land, and the point of saturation for a market. Therefore, the output equation has a certain basis in reality.

9.1.2 The Group's Output Function

Group output r is the sum total of the output of all individuals within the group. Therefore, the output function for the group is as shown in Eq. (9.2).

$$r = \sum_{i=1}^{n} r_i = \sum_{i=1}^{n} (r_0 - \frac{r_0}{1 + e_i}). \tag{9.2}$$

9.1.3 Individual's Cost Function

When n individuals exist, set the relationship between the individual's i behavioral cost c_i and behavioral effort level e_i as Eq. (9.3):

$$c_i = e_i \tag{9.3}$$

In the equation, as $e_i \geq 0,\ c_i \geq 0$.

The equation assumes that behavioral cost is proportional to effort level. For instance, in reality the time consumed during labor and other things are all proportionate with effort level (e.g., amount of work done). Hence, this equation has some basis in reality.

9.1.4 Group Cost Function

Group cost c is the sum total of the cost for all individuals within the group. Therefore, the cost function for the group is as shown in Eq. (9.4).

$$c = \sum_{i=1}^{n} c_i = \sum_{i=1}^{n} e_i \tag{9.4}$$

9.1.5 Group Utility Function

Set group utility as u and the group utility function will be Eq. (9.5):

$$u = r - c = \sum_{i=1}^{n} (r_0 - \frac{r_0}{1 + e_i}) - \sum_{i=1}^{n} e_i. \tag{9.5}$$

9.1.6 Individual Effort Levels When Group Utility Is Maximized

Assume that the individual effort level of individual i for the maximization of group utility is e_i^*. Seek the partial derivative of effort level e_i from group utility function (9.5) and set it to 0. We get:

$$\frac{\partial u}{\partial e_i} = \frac{r_0}{(1+e_i)^2} - 1 = 0$$

We then obtain the equilibrium point:

$$e_i^* = \sqrt{r_0} - 1 \tag{9.6}$$

If the objective of individual i is the maximization of group utility, then his effort level will find an equilibrium point determined in Eq. (9.6).

As we assume that there is no difference between individuals, hence when seeking the equilibrium point, each individual would put in an equal amount of effort, i.e., $e_i^* = \sqrt{r_0} - 1$ is valid for all $i \in \{1, 2, \ldots, n\}$.

In the following sections, readers will see that no matter which type of task sharing institution involved, we can conduct a comparison by using average individual effort levels and Eq. (9.6). As long as the individual's output function is consistent with Eq. (9.1) and his cost function, Eq. (9.3), when individual effort level deviates from $e_i^* = \sqrt{r_0} - 1$ in any direction, group benefits would be compromised. At the same time, as group benefits are defined as the sum total of individual benefits of all participants involved, and each individual is assumed to have the same benefit level, the compromise of group benefits would be tantamount to the compromise of individual benefits.

9.2 Institution with Reward and Cost are Both Independent: Production with No Externalities

9.2.1 The Concept of, and Sun Diagram for, the Double-Independent Institution

9.2.1.1 The Concept of the Double-Independent Institution

The double-independent institution is a basic task sharing institution. Within such an institution, each individual has exclusive enjoyment of his behavioral reward and at the same time bears his own behavioral cost entirely. This is hence known as the "double-independent" institution as the individual acts singly and autonomously in both regards.

The double-independent institution is mainly used for tasks that can be divided into individual work, such as the unloading of goods, the contracting of farm work by lots, etc.

9.2.1.2 The Sun Diagram for the Double-Independent Institution

The Sun Diagram for a double-independent institution is shown as in Fig. 9.1.

The diagram illustrates the special situation where there are two individuals in the scenario. Here, the line that leads out from e_1 points towards the positive side of the promoter (reward) r_1, indicating that the higher the effort level of individual 1 e_1, the greater r_1 will be. Similarly, the line that leads out from e_2 points towards the positive side of the promoter (reward) r_2, indicating that the higher the effort level of individual 2 e_2, the greater r_2 will be. In the figure, there is a positive correlation between only reward r_i and the individual's effort level e_i. In the same vein, there is a positive correlation between only cost c_i and the individual's effort level e_i. Hence, we see that between individuals 1 and 2 the factors of cost and reward are both mutually independent.

9.2.2 Equilibrium Model and the Equilibrium Point of Effort Level

9.2.2.1 Behavioral Reward

In the double-independent institution, the actor has exclusive enjoyment of his behavioral rewards (refers to benefits with exclusivity, if behavior type is manual labor then this type of benefit is known as labor output).

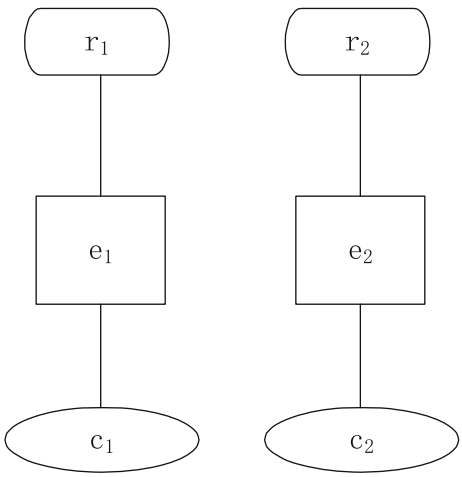

Fig. 9.1 Sun Diagram for double-independent institution (two individuals)

When n individuals exist, set the relationship function between reward r_i enjoyed by individual i and the individual's behavioral effort level e_i as Eq. (9.1) from Sect. 9.1, i.e.:

$$r_i = r_0 - \frac{r_0}{1 + e_i}$$

9.2.2.2 Behavioral Cost

When n individuals exist, set the relationship between the individual's i behavioral cost c_i and behavioral effort level e_i as Eq. (9.3):

$$c_i = e_i$$

9.2.2.3 Behavior Utility Function

When n individuals exist, set the utility function for individual i as Eq. (9.7).

$$u_i = r_i - c_i = r_0 - \frac{r_0}{1 + e_i} - e_i \tag{9.7}$$

9.2.2.4 The Equilibrium Point for Level of Effort Put into Behavior

Obtain the partial derivative of effort level from utility function (9.7) and set it to 0. We get:

$$\frac{\partial r_i}{\partial e_i} = \frac{r_0}{(1 + e_i)^2} - 1 = 0$$

We obtain the equilibrium point:

$$e_i^* = \sqrt{r_0} - 1$$

The equilibrium point is the same as the individual effort equilibrium point when group outcome is optimized [Eq. (9.6)] which we saw in the previous section. That is to say, within the double-independent institution the individual's spontaneous equilibrium point is optimal for group interests with no gap or difference between the two.

If every individual is a production enterprise, then the effort equilibrium point can be regarded as the spontaneous equilibrium point for enterprise production scale. When the individual is an enterprise, within the double-independent institution the spontaneous production scale (equilibrium) of each enterprise all contribute to the overall optimal production scale of all enterprises in society.

9.2.3 Characteristics and Applicability of the Double-Independent Institution

In contrast to the other types of the reward sharing institution to be discussed later in this chapter, in the double-independent institution the individual effort equilibrium is optimal for the group in general. Hence, the double-independent institution is the most efficient way of sharing and delegating tasks.

However, in management practice the double-independent institution cannot be used in many circumstances.

9.2.3.1 Double-Independent Institution Cannot Be Used When Individual Effort Level Cannot Be Observed Efficiently and When Group Output Cannot Be Divided to Individuals

Specifically speaking, when an output requires the coming together of multiple individuals, and when this output is not divisible (i.e., attributable to individual contributors) and when individual effort levels are not observable, the double-independent institution cannot be used. Take for instance the development of an intellectual product, such as a "new policy" by a research team. The production of the final policy involves research, conceptualization, discussions, amendments, etc., with all team members participating in every step of the process. The final "product" is hence indivisible with each individual team member's "effort level" also not observable with any precision.

On the other hand, when the discrete output of each individual can be observed (and at this point the output for each individual can be used as a basis for judging individual effort level), we can use the double-independent institution. For instance, in the case of the building of a gargantuan ship, the double-independent institution can be used for smaller components. For large jobs just as the building of a ship, for instance we can measure individual effort by measuring the area of the ship each worker has worked on. Therefore, as tasks relating to physical products are easier to divide up, the double-independent institution would be more appropriate for these cases.

Task sharing institutions where behavioral reward cannot be divided by, or attributed to individuals shall be analyzed in greater detail under the following section on the "reward-sharing institution".

9.2.3.2 Double-Independent Institution not Advisable When Behavioral Cost Cannot Be Divided

When behavioral cost is not divisible, an individual may offload his cost to other individuals so that his individual effort level exceeds the optimal point for the group. For example, when multiple fishery companies limit their "catch" to a

certain level, together they would be able to conserve fishery resources for more productive catches. However, as each company's cost—"the depletion of fishery resources"—cannot be clearly distinguished (as fish travel from zone to zone, and multiple companies often harvest fish in the same zone at the same time), each company would try to maximize their respective catches with the net result of resource depletion and less productive trips out to sea for all.

Task sharing institutions where behavioral cost cannot be divided by individuals shall be analyzed in greater detail under the following section on the "cost-pooling institution".

9.2.3.3 Performance of Double-Independent Institution Is Poor When Output Quality of Individuals Is Hard to Observer

If only the quantity and not the quality of the individual's output can be observed reliably, under the double-independent institution this can easily lead to prolific, but low-quality, production from the individual.

The double-independent institution is partially why products that are inferior yet produced in bulk like chickens and eggs are in existence.

9.3 Reward-Sharing Institutions: Production with Positive Externalities

9.3.1 The Concept of, and Sun Diagram for, the Reward-Sharing Institution

9.3.1.1 The Concept of the Reward-Sharing Institution

The reward-sharing institution is an institution that comprises multiple individuals who work together as a collective to complete a larger task. The distinctive feature of this institution is that the reward from collective output is divided equally amongst all participants. However, each individual bears his behavioral cost separately and individually. Hence, in this institution the individual shares in the group behavioral reward but bears his individual behavioral cost.

In the reward-sharing institution, if the behavior at hand is the production of goods, then this type of behavior is a kind of production with positive externalities.

In management practice, there are three main reasons for reward-sharing.

The first case is with output that brings strong positive externalities, such as the construction of highways and bridges, environment improvement efforts, etc. Apart from the producer (or investor), this type of positive externality can potentially benefit many other individuals.

The second case is with output that cannot be divided according to the number of participants in the production group. For instance, when multiple individuals pour

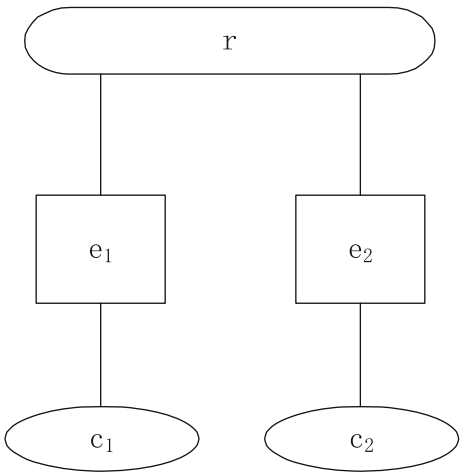

Fig. 9.2 Sun Diagram for reward-sharing institution (two individuals)

water cleaning agent into a pond, the resulting output "purified pond" is a common output that cannot be divided by the number of participating individuals.

In the third case, the institution has arisen from a result of human action, or an institution of human design. Take for instance if an equal wage institution were to be deployed with a certain group, this would lead to group output with positive externalities, that is, group behavior would enhance individual income as well as income of others, thereby creating a reward-sharing situation.

9.3.1.2 The Sun Diagram for the Reward-Sharing Institution

The Sun Diagram for a reward-sharing institution is shown as in Fig. 9.2.

In the figure, the line that leads out from e_1 as well as the line from e_2 point towards the positive side of the promoter (reward) r, indicating that the higher the effort level e_1 of individual 1, the greater r will be. The higher the effort level e_2 of individual 2, the greater r will be, too. This r is something both individuals 1 and 2 enjoy. That is to say, even when individual i puts in zero effort, as long as others put in effort, individual i would still be able to enjoy certain benefits.

9.3.2 Equilibrium Model and the Effort Equilibrium Point

9.3.2.1 Behavioral Reward

In the reward-sharing institution, each actor shares in a common pool of rewards, i.e., the reward enjoyed by each actor is the average of group rewards, i.e., total rewards divided by the number of individuals.

Have individual j's output fit into Eq. (9.1) from Sect. 9.1:

$$r_j' = r_0 - \frac{r_0}{1 + e_j}$$

When n individuals exist, the group's output is:

$$r = \sum_{j=1}^{n} r_j' = \sum_{j=1}^{n} (r_0 - \frac{r_0}{1 + e_j})$$

In the reward-sharing institution, the reward enjoyed by each individual is the average value of all rewards divided by the number of participants. Thus, the function for the relationship between reward r_i enjoyed by individual i (note that under the reward-sharing institution reward r_i enjoyed by individual i may not equal his output r_i') and group effort level is as in Eq. (9.8):

$$r_i = \frac{1}{n} \sum_{j=1}^{n} (r_0 - \frac{r_0}{1 + e_j}) \tag{9.8}$$

In the equation, $r_0 \geq 0$, $e_j \geq 0$, $j = \{1, 2, \ldots, n\}$, of which r_0 is the maximum value of the reward.

9.3.2.2 Behavioral Cost

Set the relationship between individual i's behavioral cost c_i and behavioral effort level e_i as Eq. (9.3) from Sect. 9.1:

$$c_i = e_i$$

9.3.2.3 Behavior Utility Function

Set the utility function for individual i within the reward-sharing institution as Eq. (9.9).

$$u_i = r_i - c_i = \frac{1}{n} \sum_{j=1}^{n} (r_0 - \frac{r_0}{1 + e_j}) - e_i \tag{9.9}$$

9.3.2.4 The Equilibrium Point for Level of Effort Put into Behavior

Obtain the partial derivative of effort level e_i from Eq. (9.9) and set it to 0. We get:

$$\frac{\partial u_i}{\partial e_i} = \frac{r_0}{n(1+e_i)^2} - 1 = 0$$

We then obtain the equilibrium point:

$$e_i^* = \sqrt{\frac{r_0}{n}} - 1. \tag{9.10}$$

If the individual is rational, his effort level will find an equilibrium point determined in Eq. (9.10).

9.3.3 *Characteristics of the Reward-Sharing Institution: Insufficient Individual Effort*

9.3.3.1 Effort Level Is Lower Than in the Double-Independent Institution

Compared to the effort equilibrium point $e_i^* = \sqrt{r_0} - 1$ in the double-independent institution, we see that the equilibrium point $e_i^* = \sqrt{\frac{r_0}{n}} - 1$ in the reward-sharing institution is at a lower level, as $n \geq 1$. As both group reward and individual utility are at optimal levels with the double-independent institution, group benefits and utility tend to be at a lower level with the reward-sharing institution.

If the individual is a production enterprise, its effort level is indicated by its production scale. Hence, when the enterprise's production behavior comes with positive externalities, its spontaneous production scale will stabilize at a lower point so that society's output would be at the optimal production scale (i.e., highest).

9.3.3.2 The Larger the Group, the Lower the Effort Level

As we observe the effort equilibrium level in the reward-sharing institution $e_1^* = \sqrt{\frac{r_0}{n}} - 1$, we can see that the greater the number of participants, the lower the level of individual effort. In particular, when $n = 1$, the effort equilibrium level is the same for the reward-sharing institution and the double-independent institution. When $n \to \infty$, the equilibrium point $e_i^* \to -1$ in the reward-sharing institution indicates a "negative effort level", which can be read as the individual proving to be a hindrance to others apart from not putting in any work himself.

9.3.4 Measures to Enhance Individual Effort Levels Under the Reward-Sharing Institution

Although we have no choice but to resort to the reward-sharing institution at times, we can still take certain measures to raise individual effort levels within the institution.

9.3.4.1 Decrease the Group Size as Much as Possible

As individual effort level decreases as the size of the group increases, when possible the institution designer should try to divide output such that the group size can be better managed.

9.3.4.2 The Use of High-Performance Observers

The difficulty in observing or detecting individual behavior is a key reason why the double-independent institution cannot be implemented in many cases. Therefore, the use of a high-performance observer and the granting of rewards based on observed result (for instance although each participant in the reward-sharing institution receives the same amount of wages, they can be provided bonuses of varying sizes based on their respective effort levels) in order to enhance individual effort level.

9.3.5 Case Study 1: The "Race to Be Last"

A large-scale post office is equipped with a number of counters or service windows. Each service window is manned by a member of the post office staff. In the past, the post office employed an "equal wages for all" institution with both staff wages and bonuses tied to the overall performance of the post office. The thinking behind such an arrangement was that equal wages (equal to the extent that everyone started off at a certain base rate with small increments for seniority) would keep staff content as this was a "fair institution". At the same time, staff members were also encouraged to work harder in order to improve the post office's performance (which in turn would lead to higher wages for all). This is a classic instance of the reward-sharing institution.

However, the postmaster soon found that service staff members were not enthusiastic about their work at all. During busy periods or closer to festivals when there would be a surge of customers at the post office, many service staff engaged in "a race to be last", trying ways and means to prolong the individual transaction so that they would not have to serve as many customers.

Confronted with the reality, the postmaster reformed the wage institution, tying individual wage for each month to the number of customers serviced. This is the double-independent institution at work (the behavioral cost is the effort expended by each post office employee and is independent of the effort level of the next employee). Following this reform, the "race to the last" became a thing of the past. Instead, service workers became focused on completing transactions with efficiency so that they can service more customers in a day.

9.3.6 Key Conclusions

Here, we can discover a misconception on the part of many in the worlds of economics and management science, that "when group benefits increase, benefits earned by individual participants also increase", which would then lead to "hard work among participants". From the analyses above, we see that reality does not work this way. If group output cannot be attributed to individual participants, hence making individual efforts less than observable, the individual effort equilibrium level will tend to be lower.

For instance, the "employee stock ownership plan" that is currently very popular tries to "tie employee and corporate benefits together". However, if corporate performance is determined solely by the sum total of everyone's income, then such a plan may not lead to higher individual effort levels at the workplace. Only when individual performance can be distinguished and observed, and when employees are rewarded based on their individual performance, employees can be led to work harder. Of course, shareholders with comparatively larger shareholdings will pay very close attention to the company's interests as their interests are highly correlated with those of the company's. However, for common employees who hold very few shares in comparison, the "employee stock ownership plan" may not serve as a strong motivator for work performance.

9.4 Improving the Reward-Sharing Institution: The Use of Cost Subsidies

In management practice, the reward-sharing institution is used mostly because individual output cannot be observed reliably. Hence, all participants would have to share the rewards of their work equally. However, in the reward-sharing institution there are still elements that are observable, since individual behavioral cost is an independent factor. This means that individual behavioral cost may have a higher level of observability. Hence, we can work on improving the reward-sharing institution in this area by adding cost subsidies to the reward-sharing institution to promote higher individual effort levels.

9.4.1 The Concept of, and Sun Diagram for, the Reward-Sharing Institution with Cost Subsidies Applied

9.4.1.1 Characteristics of the Reward-Sharing Institution with Cost Subsidies Applied

The reward-sharing institution with cost subsidies applied is an institution that is designed to enhance individual effort levels. This institution is similar to the regular reward-sharing institution in that rewards are shared while individual behavioral costs are independent. In addition, the (autonomous or independent) cost of the individual is subsidized to a certain degree, as cost is directly correlated with individual effort level. This means that the "cost subsidy" is also directly correlated with individual effort level. This way, the "cost subsidy" serves effectively as a kind of "reward" that enhances individual effort level. Hence, this type of subsidy can enhance individual effort level. We can imagine that individual effort level in the reward-sharing institution with cost subsidies applied would be higher than individual effort level in the pure reward-sharing institution.

We should note that the existence of the "cost subsidy" does not necessarily mean that management costs will rise, as a portion of group output can be deducted to be used as "cost subsidies". Therefore, in a reward-sharing institution with cost subsidies applied, reward per person may be slightly reduced with the amount deducted transferred to cost subsidies.

9.4.1.2 The Sun Diagram for a Reward-Sharing Institution with Cost Subsidies Applied

The Sun Diagram for a reward-sharing institution with cost subsidies applied is shown as in Fig. 9.3. Here, r_{si} is the cost subsidy targeted at individual i.

9.4.2 Equilibrium Model and the Effort Equilibrium Point

9.4.2.1 Behavioral Cost

We continue to set the relationship between individual i's behavioral cost and behavioral effort level as Eq. (9.3):

$$c_i = e_i$$

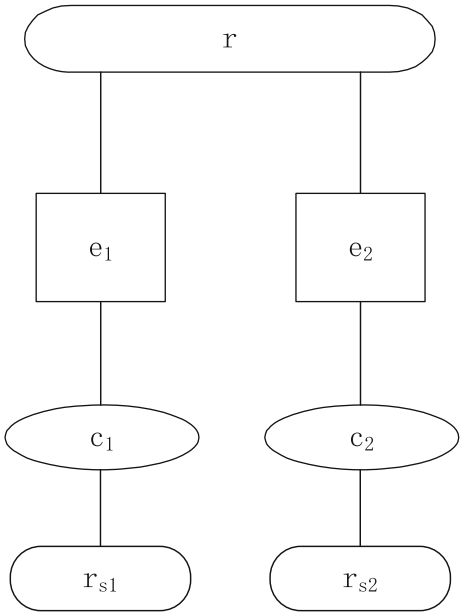

Fig. 9.3 The Sun Diagram for a reward-sharing institution with cost subsidies applied (two individuals)

9.4.2.2 Subsidy for Behavioral Cost

We can use a variety of cost functions for individuals when working on subsidies for behavioral cost. For the sake of simplicity, in this book we assume that individual i's behavioral cost subsidy r_{si} is a linear function of individual c_i. Therefore:

$$r_{si} = kc_i = ke_i \tag{9.11}$$

9.4.2.3 Behavioral Reward

In a reward-sharing institution with cost subsidies applied, reward r_i enjoyed by individual i is obtained by dividing group output (minus subsidies applied) by the number of participants in the group. Hence, the relationship between reward r_i enjoyed by individual i and group effort level is described in Eq. (9.12), i.e.:

$$r_i = \frac{1}{n}\sum_{j=1}^{n}(r_0 - \frac{r_0}{1+e_j} - r_{sj}) = \frac{1}{n}\sum_{j=1}^{n}(r_0 - \frac{r_0}{1+e_j} - ke_j) \tag{9.12}$$

9.4.2.4 Behavior Utility Function

Set the utility function for individual i within the reward-sharing institution with cost subsidies as Eq. (9.13).

$$u_i = r_i - c_i + ke_i = \frac{1}{n}\sum_{j=1}^{n}(r_0 - \frac{r_0}{1+e_j} - ke_j) - e_i + ke_i \tag{9.13}$$

9.4.2.5 The Equilibrium Point for Level of Effort Put into Behavior

Obtain the partial derivative of individual i's effort level e_i from utility function (9.13) and set it to 0. We get:

$$\frac{\partial u_i}{\partial e_i} = \frac{r_0}{n(1+e_i)^2} - \frac{k}{n} + k - 1 = \frac{r_0}{n(1+e_i)^2} + (1 - \frac{1}{n})k - 1 = 0.$$

We then derive the equilibrium point of individual i's effort level e_i:

$$e_i^* = \sqrt{\frac{r_0}{n-(n-1)k}} - 1 \tag{9.14}$$

From Eq. (9.14), we see that as $r_0 > 0$, k must meet the following conditions: $n-(n-1)k > 0$, i.e., $\frac{n}{(n-1)} > k$, at the same time from the significance of the subsidy we know that $k > 0$, hence $\frac{n}{(n-1)} > k > 0$.

Comparing Eqs. (9.14) and (9.10), as $n > n-(n-1)k$:

$$\sqrt{\frac{r_0}{n-(n-1)k}} - 1 > \sqrt{\frac{r_0}{n}} - 1.$$

Individual effort level in the reward-sharing institution with cost subsidies applied would be higher than individual effort level in the pure reward-sharing institution.

In fact, as long as group productivity is sufficiently high, i.e., the ratio of output to consumption (costs) is sufficiently high, the institution designer can choose to adjust k so that $n-(n-1)k = 1$ (i.e., k = 1, that is a full subsidy for individual behavioral cost). At this point:

$$\sqrt{\frac{r_0}{n-(n-1)k}} - 1 = \sqrt{\frac{r_0}{n-(n-1)}} - 1 = \sqrt{r_0} - 1$$

Here, the individual effort equilibrium point in the reward-sharing institution with subsidies applied is exactly the same as the individual effort equilibrium point when group benefits are maximized [i.e., Eq. (9.6)].

9.4.3 Characteristics and Applicability of the Reward-Sharing Institution with Cost Subsidies Applied

Although the use of cost subsidies in the reward-sharing institution can enhance individual effort level, there is also a flaw in this institution type: in contrast to output, cost may be less observable. Hence, using subsidies as an independent variable in the institution can lead to a false representation of actual costs.

Therefore, the reward-sharing institution with cost subsidies applied is most applicable to behaviors where cost can be observed efficiently and where there is a high correlation between behavioral cost and effort level.

9.5 The Serial-Cooperation Institution: A Variation on the Reward-Sharing Institution

9.5.1 The Concept of the Cooperation Institution

A cooperation institution is an institution where multiple individuals work in concert to form an even stronger power or force to undertake a certain task. The form of cooperation referred to here is where multiple individuals form a coordinated entity in terms of behavior, an entity that receives benefits greater than the sum total of individual benefits should participating individuals choose to "go it alone" instead. To describe this sort of situation using a mathematical model: the total output from cooperative behavior is the result of the "superadditivity" of the output of individuals.

Set r as the total group benefit for the cooperative group, $r(i)$ as individual i's individual benefit, the number of participants in the cooperative group as n. Then, define the superadditivity of $r(i)$ as $r > \sum_{i=1}^{n} r(i)$, the additivity of $r(i)$ as $r = \sum_{i=1}^{n} r(i)$, and the subadditivity of $r(i)$ as $r < \sum_{i=1}^{n} r(i)$ (such as when the cooperative group is being consumed by internal conflict). We should note that some authors have defined superadditivity as $r \geq \sum_{i=1}^{n} r(i)$. However, this definition cannot be used to distinguish superadditivity and general additivity, which is why in this book we have used $r > \sum_{i=1}^{n} r(i)$ to define superadditivity.

The cooperation institution exists as in many cases only when there is concerted effort between individuals can a task be handled effectively.

For example, when a mathematician and a physicist co-author a paper—which happens quite often—it is because the subject of the paper requires some challenging mathematical and physics work that neither the mathematician nor the physicist can complete individually. Thus, when a mathematician and a physicist co-author a paper, "superadditivity" occurs with the addition of "zero papers" on the part of the mathematician's individual output and "zero papers" from the physicist's individual output.

To take another example: the workers along all sections of a factory production line must work in order to produce the final product. No one worker can produce a product by himself. Therefore, the production volume of the line is also created by the "superadditivity" of each worker producing "zero items an hour" when working alone.

As cooperative behavior can generate a greater work capacity, the cooperative or cooperation institution is also a sub-type of the task sharing institution. At the same time, as the output of multiple individuals working in concert is often not divisible, the cooperation institution would also be classified as a reward-sharing institution in such cases.

However, in contrast to his peer in the general reward-sharing institution, in the cooperation institution an individual's output must present "superadditivity" so that the level of group benefits here can exceed that of the reward-sharing institution. In the reward-sharing institution analyzed in Sects. 9.2 and 9.3 of this chapter, individual benefits are totaled to form group benefits. For instance, when multiple individuals work together to unload goods at the speed of 20 units per person per hour, a team of 10 persons would unload at the rate of 200 units per hour. This is not the kind of cooperation that this book examines, because group output here is simply the sum total of individual output with no superadditivity involved, i.e., no "additional value" created beyond the simple sum total of individual outputs. In this scenario, even if there were only one person working to unload the goods, he would be able to complete the task anyway, albeit in a longer time frame. However, when multiple individuals are working together to lift a rock, the situation is different. If they do not cooperate, the rock cannot be lifted. In this scenario, the output of "going it alone" is zero while the benefits gained from cooperation is not merely the sum total of individual outputs but also comprises a "superadditive" component.

Cooperation institutions can be categorized as two sub-types: the serial-cooperation institution and the parallel-cooperation institution.

9.5.2 The Serial-Cooperation Institution

The term "serial-cooperation" refers to the situation where multiple individuals in a group have a multiplicative effect on group benefit levels. That is, all participants must put in effort for group benefits to happen. As long as one individual's effort level is at zero, the group benefit level would also be at zero. The size of group benefits earned is determined by the participant with the lowest effort level.

Set r as the total group benefit, the number of participants as n, e_i as the individual i's effort level, $f(\cdot)$ as the monotonically increasing function, and the mathematical definition for the multiplicative effect of cooperation as $r = f(\min(e_1, e_1, \ldots, e_n))$.

The factory production line is a classic instance of the serial-cooperation institution. The productivity of the line is determined by the slowest segment of the production and assembly process. In the most extreme scenario, when one segment of the production line breaks down, the production of the entire production line will be reduced to zero.

9.5.3 *The Sun Diagram for the Serial-Cooperation Institution*

The Sun Diagram for a serial-cooperation institution is shown as in Fig. 9.4.

Figure 9.4 shows a serial-cooperation institution with two individuals. In the serial-cooperation part of the diagram, we see individual 1 in front and individual 2 behind.

In the diagram, the arrow that originates from e_1 points towards consequence con_1, indicating that individual 1's cooperative behavior is what causes consequence con_1 to appear. The arrow that ends in consequence con_1 and the arrow that originates from e_2 both point at consequence con_2, indicating that if consequence con_1 exists, individual 2's cooperative behavior will then lead to the occurrence of

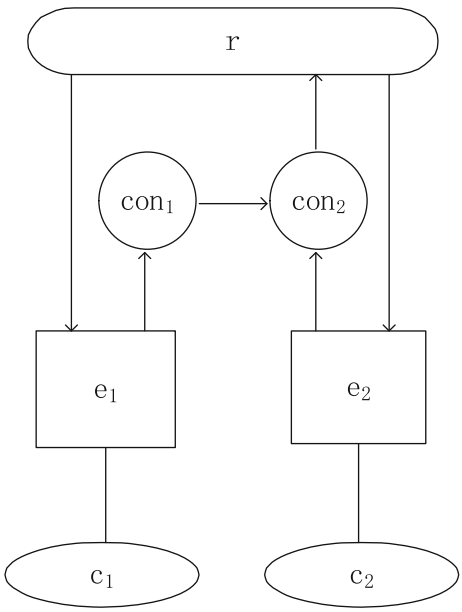

Fig. 9.4 Sun Diagram for serial-cooperation institution (two individuals and with e_1 in front)

consequence con_2. The arrow that originates from con_2 points towards shared reward r, indicating that the occurrence of con_2 will result in group benefit. An automobile production line can be seen as a case of the serial-cooperation institution. We can regard it as the preliminary assembly process in the production of the automobile. As the unit is not fully complete, no value is produced. con_2 can be regarded as the final assembly of the automobile unit which produces the ultimate form of the product. With this process, the unit acquires use value and can hence be sold on the market, generating reward r.

The two arrows that originate from the reward r point towards e_1 and e_2 respectively, indicating that the reward r is shared between individual 1 and individual 2.

We see in Fig. 9.4 that the common efforts of individuals 1 and 2 produce the shared reward r through outcomes con_1 and con_2.

Comparing Fig. 9.4 with the Sun Diagram of the pure reward-sharing institution in Fig. 9.2, we can see that in terms of institution structure the serial-cooperation institution is actually a variation on the pure reward-sharing institution. Indeed, the institution structure retains the basic feature of the reward-sharing institution, a group reward that is shared by participants, reward r. The difference between the two is, in Fig. 9.4 we have two additional serial consequences "*con*", which is a distinctive characteristic of the serial-cooperation institution.

9.5.4 *Mathematical Model for the Serial-Cooperation Institution*

9.5.4.1 Reward Function for the Serial-Cooperation Institution

(i) Reward function

Assume that there are n collaborative participants (i.e., individuals) and that benefits are distributed equally to each participant. Use Eq. (9.15) for the reward function of individual i.

$$r_i = \begin{cases} \dfrac{R_0}{n} - \dfrac{R_0}{n(1+we_i)}, & \text{当} e_i < \min\left\{\{e_1, e_2, \mathrm{K}, e_n\} \backslash \{e_i\}\right\} \\ \dfrac{R_0}{n} - \dfrac{R_0}{n(1+w(\min\{e_1, e_2, \mathrm{K}, e_n\}))}, & \text{when } e_i \geq \min\left\{\{e_1, e_2, \mathrm{K}, e_n\} \backslash \{e_i\}\right\} \end{cases} \tag{9.15}$$

In the Eq. (9.15), $R_0 > 0, n \geq 1, e_i \geq 0, w > 1$, and $\{e_1, e_2, \ldots, e_n\} \backslash \{e_i\}$ indicates that when element $\{e_i\}$ is excluded from the set $\{e_1, e_2, \ldots, e_n\}$, min $\{\{e_1, e_2, \ldots, e_n\} \backslash \{e_i\}\}$ indicates that we should seek for the element with the smallest value in the set $\{e_1, e_2, \ldots, e_n\} \backslash \{e_i\}$.

(ii) The meaning of R_0

R_0 refers to the upper limit of maximum group benefit for groups that engage in cooperative behavior. It is determined by group characteristics and cooperation specifics.

(iii) The meaning of $\min\{\{e_1, e_2, \ldots, e_n\}\backslash\{e_i\}\}$

In a group that engages in cooperative behavior, $\min\{\{e_1, e_2, \ldots, e_n\}\backslash\{e_i\}\}$ is the minimum effort level after e_i has been excluded. $e_i < \min\{\{e_1, e_2, \ldots, e_n\}\backslash\{e_i\}\}$ means that in the serial-cooperation group the member with the lowest effort level is e_i. At this point, if individual i increases his effort level e_i, group benefit $R_0 - \frac{R_0}{1+we_i}$ will increase.

$e_i \geq \min\{\{e_1, e_2, \ldots, e_n\}\backslash\{e_i\}\}$ means that in the serial-cooperation group the member with the lowest effort level is not e_i. At this point, if individual i increases his effort level e_i, group benefit $R_0 - \frac{R_0}{1+we_i}$ will not increase. Therefore, at this juncture, individual i will not put in more effort.

Hence, for individual i, $\min\{\{e_1, e_2, \ldots, e_n\}\backslash\{e_i\}\}$ determines the maximum contribution individual i makes to the group during serial cooperation (i.e., the maximum benefits realized for the group as a result of individual i's efforts).

For instance, a group of five persons (notated by the numbers 1–5) work on the same production line, on different segments and with varying effort levels. Individual 1's efforts allow him to assemble nine units every hour, while individual 2 assembles eight units every hour. Individual 3 assembles three units every hour; individual 4, seven units every hour; and individual 5, six units every hour. For individual 3, $\min\{\{e_1, e_2, e_3, e_4, e_5\}\backslash\{e_3\}\}$ would be the lowest assembly speed on the production line after excluding individual 3's speed. The lowest speed then would be individual 5's speed, i.e., $\min\{\{e_1, e_2, e_3, e_4, e_5\}\backslash\{e_3\}\}$ equals six units an hour. At this juncture, before individual 3's assembly speed reaches six units per hour, any increase in his assembly speed will enhance line productivity.

On the other hand, if individual 3's assembly speed has already reached six units per hour, any more enhancement of his assembly speed or effort level will not have any impact on group benefits. As long as individual 5's assembly speed remains at six units per hour, line production will also remain at six units per hour.

(iv) The meaning of w

w means that the cooperative group has an amplification effect on individual behavioral effect, with the size of the amplification effect related to the nature of the cooperative behavior in question. For instance, in a product institution that utilizes serial cooperation, assuming the requisite effort levels in all other segments are present, as i's effort level increases, the cooperative group's output also rises from 0. Individual i's behavioral effect is directly expressed in group output. However, as the scale of group output usually far exceeds the scale of an individual's output, individual i's output is being amplified by the cooperative group, with w being the expression for this amplification effect.

9.5.4.2 Cost Function

In a serial-cooperation scenario, individual i's behavioral cost is fully borne by the individual himself. Therefore, the cost function continues to be expressed by Eq. (9.3), i.e.:

$$c_i = e_i$$

Here, c_i stands for individual i's cooperative cost while e_i stands for individual i's effort level in cooperation.

9.5.4.3 Behavior Utility Function in the Serial-Cooperation Institution

When n individuals exist and participate in the serial-cooperation institution, set the utility function for individual i as Eq. (9.16).

$$\begin{aligned} u_i &= r_i - c_i \\ &= \begin{cases} \dfrac{R_0}{n} - \dfrac{R_0}{n(1+we_i)} - c_i, & \text{当} \ e_i < \min\left\{\left\{e_1, e_2, \mathrm{K}, e_n\right\} \backslash \left\{e_i\right\}\right\} \\ \dfrac{R_0}{n} - \dfrac{R_0}{n(1+w(\min\left\{e_1, e_2, \mathrm{K}, e_n\right\}))} - c_i, & \text{when} \ \ e_i \geq \min\left\{\left\{e_1, e_2, \mathrm{K}, e_n\right\} \backslash \left\{e_i\right\}\right\} \end{cases} \end{aligned} \tag{9.16}$$

9.5.5 Equilibrium Point of Individual Effort and Factors that Affect Effort Level During Serial Cooperation

9.5.5.1 Effort Equilibrium Point in the Serial-Cooperation Institution

Obtain the partial derivative of effort level e_i from Eq. (9.16). We get:

$$\frac{\partial u_i}{\partial e_i} = \begin{cases} \dfrac{R_0 w}{n(1+we_i)^2} - 1, & \text{当} \ e_i < \min\left\{\left\{e_1, e_2, \ldots, e_n\right\} \backslash \left\{e_i\right\}\right\} \\ -1, & \text{当} \ e_i \geq \min\left\{\left\{e_1, e_2, \ldots, e_n\right\} \backslash \left\{e_i\right\}\right\} \end{cases}$$

With $\frac{\partial u_i}{\partial e_i} = 0$, we derive individual i's effort equilibrium point:

$$e_i^* = \sqrt{\frac{R_0}{nw}} - \frac{1}{w} \tag{9.17}$$

when $e_i < \min\{\{e_1, e_2, \ldots, e_n\} \backslash \{e_i\}\}$.

9.5.5.2 Factors that Impact the Equilibrium Point in the Serial-Cooperation Institution

Let us now look at the factors that impact the equilibrium point e_i^* of individual i's effort level.

(i) The greater the number of participants, the lower the equilibrium level

From Eq. (9.17), we can see there is a positive linear relationship between e_i^* and $\frac{1}{\sqrt{n}}$. This means that the higher the number of participants n, the lower the equilibrium level e_i^* of the level of effort by the individual i.

In particular, when $n = 1$ and when $w = 1$, when there is only one individual in the cooperative group, there is no amplification effect caused by the group on individual behavioral effect. Here, Eq. (9.17) is transformed (R_0 in the equation is transformed into r_0 as in a single-individual situation maximum benefits are determined only by the nature of the individual's behavior and not by other individuals since they do not exist):

$$e_i^* = \sqrt{r_0} - 1$$

If we compare the above equation with the equilibrium equation for a single-individual institution (9.6), we will find that the two are exactly the same. That is to say, in the extreme scenario that there is only one individual within the institution, the individual effort equilibrium level in the cooperation institution will be the same as the equilibrium level under the double-independent (reward and cost-independent) institution. Hence, we see the consistency and stability of the institution classification and mathematical models in this volume.

(ii) The higher the value of R_0, the higher the equilibrium level

Now, we look at the upper limit R_0 of maximum group benefit for groups that engage in cooperative behavior. Clearly, the higher R_0 is, the higher individual i's effort equilibrium point e_i^*.

(iii) Key conclusions

From the above, we can draw some key conclusions.

First, from the analysis of the participant number n, we can see that unless there is no other choice the size of the cooperative group should be limited to maximize participant motivation and enthusiasm.

Second, from the analysis of $\min\{e_1, e_2, \ldots, e_n\} \backslash \{e_i\}$, we see that in the cooperative group the effort level of each individual is not independent of those of other individuals. The harder everyone works, the more motivated everyone is. Conversely, if no one in the group is motivated to work hard, then hard work by other individuals will also not yield the desired effect. Hence, everyone will give up on working hard altogether.

9.5.6 *The Problem of the "Wait" When Individual Effort Levels* e_k *Are Higher Than* e_i

"The problem of the "wait" when individual effort levels e_k are higher than e_i" refers to the scenario in serial cooperation where only when e_i is at its minimum level can group benefit increase as e_i increases. When e_i is at its minimum, it means that the effort level e_k of other participants k is not at minimum. This means that the part of k's effort level e_k higher than e_i's is a waste.

In such a situation, the effort level of each participant in the cooperative group can change in one of two possible ways.

First, e_i will rise continuously until the effort level of all participants reach equilibrium at a certain level. The result of this process would be Eq. (9.17).

Second, as the existence of the lowest effort level e_i means that all effort beyond e_i is a waste, all participants apart from i would lower their existing effort levels until everyone reaches equilibrium at a lower point (even at 0).

Generally speaking, the second scenario will not happen. This is because personal experience will tell all participants that if everyone maintains a certain higher effort level, the participant with the lowest effort level would also work harder for the sake of his individual benefits. This type of serial cooperation will produce comparatively more benefits, leading to everyone receiving a larger share of the rewards. However, if everyone were to engage in "a race to be last" by lowering their respective effort levels, group benefit will reach equilibrium at a very low level, leading to very little, or even zero, individual benefits for each group participant.

9.6 The Parallel-Cooperation Institution: Another Variation on the Reward-Sharing Institution

9.6.1 *The Parallel-Cooperation Institution and Its Sun Diagram*

9.6.1.1 Parallel Cooperation

The parallel-cooperation institution is where multiple individuals engage in cooperative behavior with the superadditivity effect in group benefit but not the multiplicative effect. In other words, in contrast to the serial-cooperation institution, if a single individual in the parallel-cooperation institution has an effort level of zero it would still be possible for the group to achieve some benefits.

Take for instance the lifting of a heavy object. The object cannot be lifted by an individual working alone, which is why everyone's individual effect or output here is zero. However, if everyone were to work together to lift the object, they may be able to achieve the task. The group benefit achieved here exceeds the sum total of benefits achieved by each individual. This is superadditivity at work. In such a scenario, as long as everyone chips in with enough strength, even if one participant involved is simply pretending to help and is actually not contributing any effort, the heavy object may still be lifted and the group may still achieve some benefits. This is a situation where the multiplicative effect does not exist. Here, we see that the parallel-cooperation institution is very different from the serial-cooperation institution, where the multiplicative effect is in play.

9.6.1.2 The Sun Diagram for the Parallel-Cooperation Institution

The Sun Diagram for the parallel-cooperation institution is shown as in Fig. 9.5.

Figure 9.5 shows a parallel-cooperation institution with two individuals working together. Here, the arrows that originate from e_1 and e_2 point towards consequence *con*, indicating that the cooperation between individual 1 and individual 2 will result in consequence *con*. However, consequence *con* will not have a direct impact on individual behavior e_1 and e_2. The arrows that originate from consequence *con* point towards the positive ends of the reward (promoter) *r* shared by e_1 and e_2, indicating that the appearance of *con* will lead to the augmentation of *r* while the

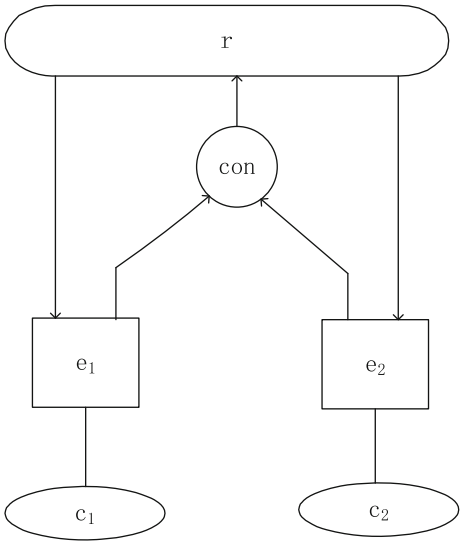

Fig. 9.5 Sun Diagram for parallel-cooperation institution (two individuals acting in concert)

arrows that originate from r point towards e_1 and e_2, indicating that individual 1 and individual 2 will enhance their effort levels e_1 and e_2 in order to obtain r.

When comparing Fig. 9.5 with the Sun Diagram of the pure reward-sharing institution in Fig. 9.2, we see that the two are highly similar. Indeed, there is no substantial difference between the two. The Sun Diagram for the pure reward-sharing institution can also be rendered as Fig. 9.5. It is only for the sake of simplicity that the consequence "*con*" was omitted from Fig. 9.2 (Sun Diagram for the reward-sharing institution). For the parallel-cooperation institution, as production in such cases often come with an effort level threshold (i.e., only when group effort reaches a certain level would production occur) or due to the significant "superadditivity" in the "consequence" that comes with cooperative behavior, we would need to emphasize the importance of consequence "*con*". Thus the Sun Diagram as shown in Fig. 9.5.

9.6.2 *Mathematical Model for Parallel-Cooperation Institution with Equal Distribution*

9.6.2.1 Reward Function for Parallel-Cooperation Institution with Equal Distribution

Assume that n participants (individuals) exist. When the individual contribution of each participant cannot be measured, the resulting benefits can only be divided equally among all participants, i.e., $r_1 = r_2 = \cdots = r_n = \frac{R}{n}$, of which R stands for group benefit. The reward function for individual i is shown in Eq. (9.18).

$$r_i = \begin{cases} \dfrac{R_0}{n} - \dfrac{R_0}{n(1+we_i)}, & \text{当} \sum e_j \geq e_0 \\ 0, \quad \text{when } \sum e_j < e_0 \end{cases} \tag{9.18}$$

In the equation, $R_0 > 0, n \geq 1, e_i \geq 0, w > 0, e_0 > 0$.

The reward function for cooperation (9.18) reflects the fact that when $\sum e_j \geq e_0$, the reward enjoyed by individual i increases as his effort level e_i rises. However, the degree of increase will taper off as e_i increases.

The definitions of various key elements in Eq. (9.18) are as follows.

(i) The meaning of R_0

R_0 means that the upper limit of maximum group benefit when the individual raises his effort level (as determined by the specific nature of the cooperative behavior).

(ii) The meaning of e_0

e_0 means that the minimum "cooperative effort" needed to generate group benefits, i.e., the minimum sum total of individual participants' effort levels needed to generate group benefit (note that this sum can comprise a super-additive component). This is related to the nature of the cooperative behavior in question. For instance, when a 300 kg object needs to be transported, only with the work of multiple individuals can the object be lifted (i.e., group benefit). Hence, $e_0 = 300$ (if everyone's effort level is measured by their strength used in lifting).

(iii) The meaning of w

w is the amplification factor for individual behavior e_i (determined by specific nature of cooperative behavior) during cooperation in a parallel-cooperation institution. Under ordinary circumstances, $w > 1$, which indicates that the cooperative group has an amplifying effect on individual behavior. For example, if five persons work together to lift an object that weighs 300 kg, if the combined lifting power of the first four persons is 240 kg then the last individual would need to put in 60 kg of lifting power to move the (300 kg) object. In this case, the amplification factor of w on e_i is $w = \frac{300}{60}$.

If $w = 1$, then it means that the cooperative group has neither an amplification or diminishing effect on individual behavioral benefits, with group benefits only the pure sum total of individual benefits. Here, there is no superadditive to the benefits earned by individual members of the cooperative.

When $w < 1$, it means that the cooperative does not have an amplification effect on individual behavioral benefits; instead, the former has a diminishing effect on the other. That is, there is no "superadditivity" between individual benefits although "subadditivity" exists in this case. At this point, the individual's contribution to group benefits is lower than if he would achieve by "going it alone". In reality, such cooperative groups do exist, such as in the case of groups consumed by internal conflict.

9.6.2.2 Cost Function for the Serial-Cooperation Institution

In a parallel cooperation scenario, the individual i's cost function is still assumed to be Eq. (9.3), i.e.:

$$c_i = e_i$$

Here, c_i stands for the individual i's cooperative cost while e_i stands for the individual i's effort level in cooperation.

9.6.2.3 Behavior Utility Function in the Serial-Cooperation Institution

With equal distribution, the utility function for individual i is as in Eq. (9.19).

$$u_i = r_i - e_i = \begin{cases} \dfrac{R_0}{n} - \dfrac{R_0}{n(1+we_i)} - e_i, \text{当} \sum e_j \geq e_0 \\ 0 - e_i, \quad \text{when } \sum e_j < e_0 \end{cases}$$

i.e.:

$$u_i = \begin{cases} \dfrac{R_0}{n} - \dfrac{R_0}{n(1+we_i)} - e_i, \text{当} \sum e_j \geq e_0 \\ 0 - e_i, \quad \text{when } \sum e_j < e_0 \end{cases} \tag{9.19}$$

9.6.3 *Individual Effort Equilibrium Point and Factors that Affect Its Level in a Parallel-Cooperation Institution with Equal Distribution*

Obtain the partial derivative of effort level e_i from Eq. (9.19). We get:

$$\frac{\partial u_i}{\partial e_i} = \begin{cases} \dfrac{R_0 w}{n(1+we_i)^2} - 1, \text{当} \sum e_j \geq e_0 \\ -1, \quad \text{when } \sum e_j < e_0 \end{cases}$$

With $\frac{\partial u_i}{\partial e_i} = 0$, we derive individual i's effort equilibrium point:

$$e_i^* = \sqrt{\frac{R_0}{nw}} - \frac{1}{w} \quad \text{when} \sum e_j \geq e_0 \tag{9.20}$$

From Eq. (9.20), we see that the individual effort equilibrium point in the parallel-cooperation institution (where benefits are distributed equally) is exactly the same as the individual effort equilibrium point in the serial-cooperation institution [where benefits are also distributed equally, Eq. (9.17)].

Hence, the factors that affect the equilibrium point in the parallel-cooperation institution are the same as those for the serial-cooperation institution. That is, the larger the group size, the lower the equilibrium level. The higher the group benefit limit R_0 in the cooperative group, the higher the equilibrium level.

9.7 The Cost-Pooling Institution: Production with Negative Externalities

9.7.1 The Concept of and Sun Diagram for the Cost-Pooling Institution

9.7.1.1 The Concept of the Cost-Pooling Institution

The cost-pooling institution is also an institution where multiple individuals work together to complete a large task. The distinctive feature of this institution is that while the cost of productive behavior is shared equally by all individuals, each individual has exclusive enjoyment of the reward granted to him.

If the behavior in question is productive behavior, then we are looking at production with negative externalities.

There are two main causes of cost-pooling.

The first cause is an institution one: cost-pooling caused by management institution structures. Take for instance the case of a taxi company that has a shared fuel station for drivers. Each driver may top up his tank to whatever level he wishes, with the amount of fuel drawn not recorded. Hence, the company only knows how much fuel is taken in total, with this figure then divided equally between all drivers. On the other hand, each driver is entitled to his earnings, and the more he works, the more he earns. Take another example, this time of a construction company which has a supply of tools such as shovels and picks for the use of its workers. Here, the cost of the equipment is divided equally among all workers. However, workers' outputs are individually evaluated and hence individual output is autonomous. The same applies here: the more the worker works, the more he earns.

The second cost is that of externality-pooling. This is a type of cost-pooling caused by the inherent nature of the behavior in question. For example, when a production process causes environmental pollution, while the benefits of such production go directly to the enterprise, the local government and people have to bear with the resulting pollution. On the other hand, while the operating income of a logistics company goes directly to the company, the traffic congestion caused by its transport behavior has to be borne by the society in general. This type of negative externalities chiefly arises due to the indivisibility of certain labor costs, in particular resource consumption. Take for instance offshore fishing: as schools of fish travel freely from zone to zone, we cannot attribute resource depletion to a specific group of fishermen. However, each fishing boat's catch (i.e., reward) is enjoyed solely by the owner and crew of the boat.

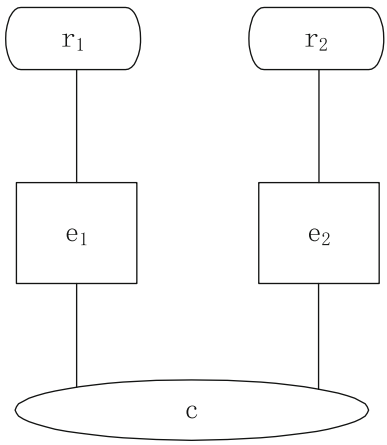

Fig. 9.6 Sun Diagram for a cost-pooling institution (two individuals)

9.7.1.2 The Sun Diagram for the Cost-Pooling Institution

The Sun Diagram for a cost-pooling institution is shown as in Fig. 9.6.

In the figure, the line that leads out from e_1 and the line from e_2 point towards the positive side of cost c, indicating that the higher the effort level e_1 of individual 1, the greater c will be. The higher the effort level e_2 of individual 2, the greater c will be, too. This cost c is shared between individuals 1 and 2. That is to say, even when individual i does not engage in any production behavior, as long as others engage in such behavior the individual would need to bear a certain cost.

9.7.2 Equilibrium Model and Effort Equilibrium Point

9.7.2.1 Behavioral Reward

In the cost-pooling institution, each actor enjoys an independent reward. Hence, just like with the double-independent institution, the relationship between individual i's output r_i and his effort level e_i is expressed in Eq. (9.1):

$$r_i = r_0 - \frac{r_0}{1 + e_i}$$

9.7.2.2 Behavioral Cost

(Average) behavioral cost is the sum total of all costs incurred by participants in the course of behavior and then divided by the number of total participants. As such, set

the relationship between the individual's i behavioral cost and the group's behavioral effort level as Eq. (9.21):

$$c_i = \frac{1}{n}\sum_{j=1}^{n} e_j \tag{9.21}$$

In the equation, $e_j \geq 0$, $j = \{1, 2, \ldots, n\}$.

9.7.2.3 Behavior Utility Function

Set the utility function for individual i within the cost-pooling institution as Eq. (9.22).

$$u_i = r_i - c_i = r_0 - \frac{r_0}{1 + e_i} - \frac{1}{n}\sum_{j=1}^{n} e_j \tag{9.22}$$

9.7.2.4 The Equilibrium Point for Level of Effort Put into Behavior

Obtain the partial derivative of effort level e_i from utility function (9.22), and set it to 0. We get:

$$\frac{\partial u_i}{\partial e_i} = \frac{r_0}{(1 + e_i)^2} - \frac{1}{n} = 0$$

We then derive the equilibrium point of individual i's effort level e_i:

$$e_i^* = \sqrt{nr_0} - 1 \tag{9.23}$$

If the individual i is a rational individual, then his effort level e_i will find an equilibrium point determined in Eq. (9.23).

9.7.3 Characteristics of the Cost-Pooling Institution: Excess Individual Effort

9.7.3.1 Effort Level Is Comparatively Higher Than in the Double-Independent Institution

Compared to the effort equilibrium point $e_i^* = \sqrt{r_0} - 1$ in the double-independent institution, we see the equilibrium point $e_i^* = \sqrt{nr_0} - 1$ in the cost-pooling institution is at a comparatively higher level, as $n \geq 1$. As both group reward and group

utility are at optimal levels with the double-independent institution, this means that individual effort level has deviated from the equilibrium point, making group benefits and utility tend lower with the cost-pooling institution.

Here, we must specify the meaning of "effort level [tending] higher". In production activity and the like, "raising effort level" means the expansion of production scale and also implies a greater speed at which production resources are being consumed. On the other hand, "cost pooling" means that production resources are shared. For instance, in some institutions the fish in marine fisheries are shared, as with the trees in forests, the grasslands for nomads, etc. Since "reward is independent" in the cost-pooling institution, it means that the individual has sole enjoyment of his production output. Hence, the fishing boat has sole enjoyment of its case, the lumberjack his timber, and the nomad the wool from his sheep bred on shared grasslands.

Hence, in the cost-pooling institution, individual effort levels will tend higher, meaning that resources would be depleted. As a result, production efficiency will be diminished. In the case of the marine catch, the quantity of fish caught would decrease significantly as a result of over-fishing, making each journey out to sea less productive; over-deforestation would cause the number of trees in the forest to decrease rapidly, making for less efficient timber harvesting; the depletion of grasslands by nomads has also led to grasslands of poorer quality, etc. Clearly, effort levels tending high have a negative impact on productivity and hence group benefit levels, and lower individual benefit levels received as well.

When the individual is an enterprise, the above analysis shows that when production activity presents negative externalities, the spontaneous production scale of the enterprises would tend greater, leading to the depletion of production resources and lower enterprise productivity as a result. This would in turn lead to lower benefit levels for the whole of society as well as for the enterprise.

9.7.3.2 The Larger the Group, the Higher the Effort Level

As we observe the effort equilibrium level in the cost-pooling institution $e_i^* = \sqrt{nr_0} - 1$, we can see that the greater the number of participants, the higher the level of individual effort. In particular, when the number of participants in the group decreases to the minimum, i.e., $n = 1$, the effort equilibrium level is the same for the reward-sharing institution and the double-independent institution. When $n \to \infty$, the equilibrium point $e_i^* \to \infty$ in the reward-sharing institution can be understood as the full exhaustion of production resources, with individual reward consequently at 0.

9.7.3.3 Applicability of the Cost-Pooling Institution

From the perspective of resource conservation, the cost-pooling institution is clearly inferior to the double-independent institution. However, in practice as we often cannot observe individual output levels, the cost-pooling institution has to be used

instead. For example, the harvesting of wild herbs [for sale] is in effect a cost-pooling institution. Under such circumstances, if we cannot improve the cost-pooling institution (we will look at how the cost-pooling institution can be improved in the following section) the number of wild herb plants will continue to decrease and the productivity of harvesters will be affected as a result. The story of the valuable herb, cordyceps, is an excellent illustration of this principle.

9.7.4 Measures to Lower Individual Effort Levels Under the Cost-Pooling Institution

Although we have no choice but to use the cost-pooling institution at times, we can take some measures to lower individual effort levels so that they deviate less from the optimal group level. This includes limiting the group size as much as possible. This is because based on the effort equilibrium equation for the cost-pooling institution, Eq. (9.23), the bigger the group, the higher the individual effort level will tend. Hence, as much as possible the institution designer should try to limit the size of the group so that individual effort equilibrium levels deviate less from the optimal group level.

9.8 Improving the Cost-Pooling Institution: Application of Tax Collections

The cost-pooling institution is arisen in management practice mainly in situations where behavioral cost cannot be divided among group participants. However, there are observable elements within a cost-pooling institution, such as individuals' behavioral reward (i.e., individual output when the target behavior is "production") as individual reward is autonomous. We can make certain improvements to the cost-pooling institution, such as tax collection, to lower individual effort levels in order to regulate resource consumption and to ensure that institution equilibrium moves closer to the optimal level.

9.8.1 The Concept of and Sun Diagram for the Cost-Pooling Institution with Tax Collections

9.8.1.1 Characteristics of the Cost-Pooling Institution with Tax Collections

The cost-pooling institution with tax collections applied is designed to lower individual effort level so as to conserve resources. Further, this institution

equilibrium can help to ensure that institution equilibrium moves closer to the optimal level.

Apart from the characteristics of cost-sharing and autonomous rewards, this institution also imposes a certain "tax" on individual reward, a tax that is generally collected with the scale of the individual's reward used as basis for calculation. As individual reward is autonomous, each individual's reward is only determined by his effort level. This means that any "tax" collected from the individual will have a direct impact on the individual's effort level. Hence, "taxes" here effectively work as a "negative reward" or "penalty" for individuals' raised effort level. Hence, tax collection has the effect of lowering individual effort levels. We can imagine that individual effort level in the cost-pooling institution with taxes applied would be lower than individual effort level in the pure cost-pooling institution.

9.8.1.2 Sun Diagram for the Cost-Pooling Institution with Tax Collections

The Sun Diagram for a cost-pooling institution with tax collections is shown as in Fig. 9.7. Here, s_{ti} is the tax collection targeted at individual i.

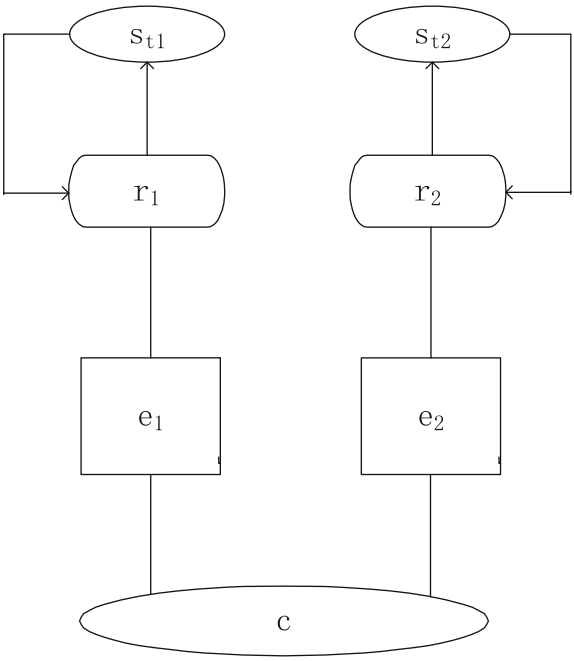

Fig. 9.7 Sun Diagram for a cost-pooling institution with tax collections (two individuals)

9.8.2 Equilibrium Model and the Effort Equilibrium Point

9.8.2.1 Tax Collections Targeted at Individual Output

For tax collection, we can use various functional equations for individual output. For the sake of simplicity, in this volume we assume tax collection to be a linear function of individual output, i.e.:

$$s_{ti} = hr_i' \tag{9.24}$$

$r_i' \geq 0$ is the output of individual i. $i = \{1, 2, \ldots, n\}$. $1 > h > 0$, h is called the tax rate.

Then, use Eq. (9.1) for the output function of individual i:

$$r_i' = r_0 - \frac{r_0}{1 + e_i}$$

9.8.2.2 Behavioral Reward for the Individual

In the cost-pooling institution with tax collections applied, the reward enjoyed by individual i is what remains after tax is deducted from individual output. Hence, the functional equation that describes the relationship between reward r_i enjoyed by individual i and behavioral effort level e_i is Eq. (9.25):

$$r_i = r_i' - hr_i' = (1 - h)r_i' = (1 - h)(r_0 - \frac{r_0}{1 + e_i}) \tag{9.25}$$

In the equation, $r_0 \geq 0, e_i \geq 0, i = \{1, 2, \ldots, n\}$, of which r_0 is the maximum output level.

9.8.2.3 Behavioral Cost for the Individual

As the cost-pooling institution with tax collections applied is still a cost-pooling institution, individual i's behavioral cost function remains as Eq. (9.21):

$$c_i = \frac{1}{n}\sum_{j=1}^{n} e_j$$

In the equation, $e_j \geq 0, j = \{1, 2, \ldots, n\}$.

9.8.2.4 Behavioral Utility Function for the Individual

Set the utility function for individual i within the cost-pooling institution with tax collections as Eq. (9.26).

$$u_i = r_i - c_i = (1-h)(r_0 - \frac{r_0}{1+e_i}) - \frac{1}{n}\sum_{j=1}^{n} e_j \tag{9.26}$$

9.8.2.5 Equilibrium Point for Level of Effort Put into Behavior

Obtain the partial derivative of effort level e_i from Eq. (9.26) and set it to 0. We get:

$$\frac{\partial u_i}{\partial e_i} = \frac{r_0(1-h)}{(1+e_i)^2} - \frac{1}{n} = 0$$

We then derive the equilibrium point of individual i's effort level e_i:

$$e_i^* = \sqrt{nr_0(1-h)} - 1 \tag{9.27}$$

If the individual is a rational individual, his effort level will find an equilibrium point determined in Eq. (9.27).

Comparing Eqs. (9.27) and (9.23), as $1 > h > 0$, $1 > 1 - h > 0$, hence:

$$\sqrt{nr_0} - 1 > \sqrt{nr_0(1-h)} - 1$$

Thus, when the individual's effort level reaches equilibrium, with the tax collection component in the cost-pooling institution individual effort level would be lower than individual effort level in the basic cost-pooling institution.

In reality, the institution designer can also adjust tax rate h to achieve $1 - h = \frac{1}{n}$. At this point, the individual effort equilibrium point in the cost-pooling institution with tax collection applied is:

$$e_i^* = \sqrt{nr_0(1-h)} - 1 = \sqrt{nr_0 \times \frac{1}{n}} - 1 = \sqrt{r_0} - 1$$

That is, the individual effort equilibrium point in the cost-pooling institution with tax collection applied is exactly the same as the individual effort equilibrium point when group benefits are maximized [i.e., Eq. (9.6)].

In addition, in the cost-pooling institution with tax collection applied, from the common condition $1 - h = \frac{1}{n}$ for the individual effort equilibrium point and the optimal group equilibrium point, we can see that to align the two, when greater the

group size, i.e., the greater n is, the higher the tax rate h has to be. In particular, when $n \rightarrow \infty$, tax rate $h \rightarrow 1$, i.e., almost all individuals are required to pay taxes on their output.

9.8.3 Key Conclusion: Significance of Tax Collection

From the analyses above, we can see that the significance of tax collection is not just limited to the redistribution of wealth as taught in classical economics. Here, tax collection can be used as a tool to regulate production speed as part of efforts to conserve more vulnerable (but renewable) resources such as fisheries, fauna and flora so that they would not be completely depleted. While the natural environment has the capacity to absorb the impact of pollution to a certain degree, once past this level this capacity will no longer be in use. Here, the environment is also a vulnerable renewable resource, and therefore tax collection may contribute to the important cause that is environmental protection.

Hence, the principles explained in this chapter can help to explain why some countries with high taxation rates have been able to keep their environments clean and eco-institution protected.

9.9 The Composite Task Sharing Institution

9.9.1 The Complexity of Reality

The three types of task sharing institutions described above are all absolute expressions of real-life situations from a management practice perspective. However, real-life situations are often in "in-between zones".

9.9.1.1 The Co-existence of Reward Sharing and Reward Independent

For instance, looking at behavioral reward, institutions often have both independent and shared reward components deployed.

An example would be the case of a certain working unit that has implemented a "points institution" for employee performance. Under this institution, each employee would receive a specific year-end bonus depending on the number of points he has earned over the year. On the other hand, the "point value" for each unit is determined by the overall performance of the unit. When the unit has turned in a strong performance, each "point" would be worth more, for example at "RMB20,000 per point". Conversely, when unit performance is poor, each "point"

may be worth less in monetary value, such as at "RMB10,000 per point". This way, an employee who has scored five points over the year would receive RMB100,000 in bonuses if the unit has performed well, and only RMB50,000 if unit performance is found to be lacking. Here, we see that with this institution the reward that the employee receives has an independent component in that the reward is partly determined by individual performance (as expressed in points). However, how much a point is worth in monetary terms is tied to overall unit performance.

9.9.1.2 The Co-existence of Cost Independent and Cost Sharing

It is the same situation with behavioral cost. In the individual's behavioral cost, very often we can find both an independent and shared component. For example, the physical and mental efforts put in during labor are considered to be autonomous costs. While the equipment used, venue and roads, environmental pollution caused (noise, air, water, land, etc.), and other impacts that would deprive others of production opportunities and cause inconvenience to others are considered shared costs.

Hence, in real life the task sharing institution can be more complex and can be an amalgam of more than one institution type.

9.9.2 The Composite Reward and Cost Functions

9.9.2.1 The Composite Reward Function

The composite reward function comprises two parts: the independent reward and the shared reward.

(i) The function for independent reward r_i'' in individual i's reward is the same as the equation for individual reward when group benefits are at the optimal level, i.e., Eq. (9.1):

$$r_i'' = r_0 - \frac{r_0}{1 + e_i}$$

(ii) The reward function for the shared average reward r_i''' in individual i's reward is the same as the equation for the reward received in the reward-sharing institution, i.e., Eq. (9.8):

$$r_i''' = \frac{1}{n}\sum_{j=1}^{n}\left(r_0 - \frac{r_0}{1 + e_j}\right)$$

(iii) The composite reward function for individual i

Set $0 \leq \alpha \leq 1$ as the scaling factor for the individual's independent reward component in a composite task sharing institution, which will render the composite reward function for reward r_i to individual i as:

$$r_i = \alpha r_i'' + (1-\alpha) r_i''' = \alpha(r_0 - \frac{r_0}{1+e_i}) + \frac{(1-\alpha)}{n} \sum_{j=1}^{n} (r_0 - \frac{r_0}{1+e_j}) \tag{9.28}$$

From the equation we can see that when $\alpha = 1$ the composite reward function has become a pure independent reward, which makes it the same as in Eq. (9.1):

$$r_i = r_0 - \frac{r_0}{1+e_i}$$

When $\alpha = 0$, the composite reward function has become a pure shared reward, which makes it the same as in Eq. (9.8):

$$r_i = \frac{1}{n} \sum_{j=1}^{n} (r_0 - \frac{r_0}{1+e_j})$$

9.9.2.2 The Composite Cost Function

The composite cost function for individual i comprises two parts: a shared cost and an independent cost.

(i) The function for independent cost in individual i's cost is the same as the equation for individual cost when group benefits are at the optimal level, i.e., Eq. (9.3):

$$c_i' = e_i$$

(ii) The function for the portion of the shared cost borne by individual i in the cost-pooling institution is the same as the cost function in the cost-pooling institution, i.e., Eq. (9.21):

$$c_i'' = \frac{1}{n} \sum_{j=1}^{n} e_j$$

(iii) The composite cost function for individual i

Set $0 \leq \beta \leq 1$ as the scaling factor for the individual's independent cost component in a composite task sharing institution, which will render the composite cost function for the cost c_i to individual i as Eq. (9.29):

$$c_i = \beta c_i' + (1-\beta)c_i'' = \beta e_i + \frac{1-\beta}{n}\sum_{j=1}^{n} e_j \tag{9.29}$$

From the equation we can see that when $\beta = 1$ the composite cost function has become a pure independent cost, which makes it the same as in Eq. (9.3):

$$c_i = e_i$$

When $\beta = 0$, the composite cost function has become a pure pooled cost, which makes it the same as in Eq. (9.21):

$$c_i'' = \frac{1}{n}\sum_{j=1}^{n} e_j$$

9.9.3 *Individual Utility Function and Effort Equilibrium Point Under the Composite Task Sharing Institution*

The function for individual utility u_i in a composite task sharing institution is laid out in Eq. (9.30):

$$\begin{aligned} u_i &= r_i - c_i \\ &= \alpha(r_0 - \frac{r_0}{1+e_i}) + \frac{(1-\alpha)}{n}\sum_{j=1}^{n}(r_0 - \frac{r_0}{1+e_j}) - \beta e_i - \frac{1-\beta}{n}\sum_{j=1}^{n} e_j \end{aligned} \tag{9.30}$$

Obtain the partial derivative of effort level e_i from Eq. (9.30) and set it to 0. We get:

$$\frac{\partial u_i}{\partial e_i} = \frac{\alpha r_0}{(1+e_i)^2} + \frac{(1-\alpha)r_0}{n(1+e_i)^2} - \beta - \frac{1-\beta}{n} = 0$$

We then derive the equilibrium point of individual i's effort level e_i:

$$e_i^* = \sqrt{\frac{[(n-1)\alpha + 1]r_0}{(n-1)\beta + 1}} - 1 \tag{9.31}$$

9.9.4 The Transformative Relationship Between the Composite Task Sharing Institution and the Three Basic Institution Types

When the scaling factor α for independent reward and the scaling factor β for independent cost take on extreme values, the composite task sharing institution then becomes a basic task sharing institution.

9.9.4.1 Double-Independent Institution When Both Scaling Factors Are 1

In Eq. (9.31), set:

$$\begin{cases} \alpha = 1 \\ \beta = 1 \end{cases}$$

While Eq. (9.31) is transformed into:

$$e_i^* = \sqrt{\frac{[(n-1)\alpha + 1]r_0}{(n-1)\beta + 1}} - 1 = \sqrt{\frac{[(n-1)+1]r_0}{(n-1)+1}} - 1 = \sqrt{r_0} - 1$$

That is, Eq. (9.31) is transformed into the effort equilibrium Eq. (9.6) for optimal individual effort level in the double-independent institution.

9.9.4.2 Reward-Sharing Institution When α is 0 and β is 1

In Eq. (9.31), set:

$$\begin{cases} \alpha = 0 \\ \beta = 1 \end{cases}$$

While Eq. (9.31) is transformed into:

$$e_i^* = \sqrt{\frac{[(n-1)\alpha + 1]r_0}{(n-1)\beta + 1}} - 1 = \sqrt{\frac{[0+1]r_0}{(n-1)+1}} - 1 = \sqrt{\frac{r_0}{n}} - 1$$

That is, Eq. (9.31) is transformed into the effort equilibrium Eq. (9.10) for optimal individual effort level in the reward-sharing institution.

9.9.4.3 Cost-Pooling Institution When α is 1 and β is 0

In Eq. (9.31), set:

$$\begin{cases} \alpha = 1 \\ \beta = 0 \end{cases}$$

While Eq. (9.31) is transformed into:

$$e_i^* = \sqrt{\frac{[(n-1)\alpha + 1]r_0}{(n-1)\beta + 1}} - 1 = \sqrt{\frac{[(n-1)+1]r_0}{0+1}} - 1 = \sqrt{nr_0} - 1$$

That is, Eq. (9.31) is transformed into the effort equilibrium Eq. (9.23) for optimal individual effort level in the cost-pooling institution.

9.9.5 Chapter Conclusions

From the discussions in this chapter, we can see that there are three basic types of the task sharing institution, i.e., the double-independent institution (the management institution for production behavior with no externalities), the reward-sharing institution (the management institution for production behavior with positive externalities), and the cost-pooling institution (the management institution for production behavior with negative externalities). Of these three, the individual effort equilibrium point under the double-independent institution is the most optimal (when the individual is an enterprise, it means that the firm's spontaneous production scale is most appropriate), while individual effort levels tend to be lower in the reward-sharing institution (when the individual is an enterprise, it means that its spontaneous production scale is smaller than optimal). With the cost-pooling institution, individual effort level tends to be high (when the individual is an enterprise, its spontaneous production scale is larger than optimal).

Based on the analyses on this chapter, we have come to the following six key conclusions with regard to institution design:

(i) When behavioral cost and behavioral reward can be split between individuals and when there are zero externalities from production behavior, we should first consider the double-independent institution (that is, the free-market model where profits and losses are borne solely by operators) as under this institution individual effort levels will spontaneously reach their optimal levels (when the individual is an enterprise, this would mean that it would spontaneously reach its optimal production scale) without the need for management costs.

(ii) When behavioral cost cannot be split between individuals and when production presents positive externalities, the production scale of the enterprise will have a lower spontaneous equilibrium point, which means that the

enterprise cannot reach its optimal scale of production. One key finding of this chapter is that the subsidy mechanism can lead to the spontaneous growth of production scale where production scale has positive externalities. Therefore, if we want the enterprise to spontaneously grow its production scale, we would need to adopt a cost subsidy policy for the enterprise. Production with positive externalities includes production activity like the construction of basic infrastructure such as reservoirs, roads and bridges.

(iii) When behavioral cost cannot be split between individuals and when production presents negative externalities, the production scale of the enterprise will have a higher spontaneous equilibrium point, leading to depletion of resources. With regard to this issue, the result of the analysis contained in this chapter shows that tax mechanisms can be used to shrink the scale of production with negative externalities so that resources may be better conserved. Hence, a high-tax mechanism needs to be put in place to shrink enterprise production scale in order to better protect the environment and to conserve resources. Such production activities come with negative externalities such as the consumption of vast amounts of water, mineral resources, fishery and forestry resources, etc., or atmospheric, water, and soil pollution. One discovery made here is that the importance of tax collection lies not in the re-distribution of citizens' income but in reducing the over-consumption of various resources. This explains the puzzle of why countries with high tax rates tend to have better-protected natural resources and environments.

(iv) Therefore, from the perspective of management by the state, when production behavior presents either positive or negative externalities, either the free-market mechanism of independent operation by each enterprise be abandoned in favor of a centralized management mechanism or the free-market mechanism be retained with supplementation by policies.

The significance of this conclusion is that it has proved, in theory, that the key approach to improving the supply of public goods and the protection of our earth lies in the institution.

Around the world, in some countries production behavior that presents externalities is put under centralized management while in others, the free market is allowed to operate with regulatory policies in place. However, with the growing scale and increasing complexity of social institutions, centralized management is becoming a less attractive option by the day for reasons of the need to communicate across multiple players/levels, significant distortion and loss of fidelity as orders and instructions percolate through a hierarchy. In a management institution where the free market is supplemented with regulatory policies, the information chain with an autonomously-run enterprise is much shorter, which means that the enterprise is more responsive. If the state implements its regulatory policies effectively, the production scale of the enterprise will grow swiftly with self-regulation and reach an equilibrium point with optimal social benefits. Relatively speaking, this would mean a more successful outcome.

(v) For some enterprises, different externalities associated with production emerge under various circumstances. For instance, for enterprises in the cultural industry, such as those in publishing and journalism, if their products drive the development of society (such as when they publish motivational tomes or volumes with scientific content), their production comes with positive externalities. If their products bring harm to society, such as when their products promote superstition, contain pornographic content, etc., their production presents negative externalities. Based on the analysis in this chapter, if these enterprises are allowed to run their operations with complete freedom as in the free-market model, positive externalities associated with their production will lead to a smaller production scale due to lack of effort while on the other hand production with negative externalities will tend to become greater as a higher level of effort is applied. Therefore, for enterprises of this type, either centralized management should be applied across the board or governmental subsidies should be applied to production that presents positive externalities to promote the scale of such production while taxes shall be levied on production that presents negative externalities to reduce the profits that can be gained from such production.

(vi) For research and development ("R&D") behavior which inherently presents positive externalities, as R&D outcomes can easily be copied (with an externality produced at this juncture, i.e., benefits for the copier), the equilibrium point of self-motivated effort is comparatively lower. Hence, currently in various countries there are measures that protect innovators, i.e., patent rights measures. This type of institution can lead to a lack of visible externalities from R&D behavior within a certain period (with the copying of R&D results in this period being illegal), thereby optimizing the effort equilibrium for R&D behavior. The problem is that if patent rights institutions are not run effectively and copying behavior cannot be hence curbed effectively, individuals would then lose their enthusiasm for R&D and innovation.

Chapter 10
Benefit Distribution Institution: Competitive Institution

Benefit distribution institutions is designed to give out "benefits" to each and every individual within an institution. In social life, many "benefits" are often limited and there is not enough to go around for everyone. Therefore, the benefits available must be distributed to individuals in accordance with certain standards or principles to maximize the effect of these benefits in society. For example, the opportunities to access higher education, to get jobs and promotions, to win a greater market share for products all are distributed to individuals through a benefit distribution institution, from an institution management perspective.

There is fundamentally only one type of benefit distribution institution, the competitive institution.

10.1 Competitive Institution

10.1.1 About the Competitive Institution

10.1.1.1 The Concept of the Competitive Institution

Competition is the targeted behavior aimed at seeking benefits, and through which individuals seek to eliminate the adverse effects posed by other individuals. We need to note the difference here between competition and opposition: the latter is when the individual in question has a goal opposite to that of his counterpart(s), such as when A says X, B disagrees. The fundamental feature of a competitive institution is that competitors engage in the behavior that helps them to "surpass other competitors".

The competitive institution is a basic benefit distribution institution where in a group made up of individuals under management, the behavioral reward for each individual increases with effort level on the part of the self and decreases with effort level on the part of other individuals.

The competitive institution is used mainly for the distribution of limited benefits of all sorts in order to maximize the social effect of these benefits.

S. Sun, *Five Basic Institution Structures and Institutional Economics*,
DOI 10.1007/978-981-10-0344-8_10

In a competitive institution, the benefit to be distributed is usually set as the object of competition, which competitors then compete for through the increase of the effort levels of their competitive behavior.

10.1.1.2 Two Functions of the Competitive Institution: Benefit Distribution and Behavior Promotion

(i) The foremost function of the competitive institution is benefit distribution.

With regard to the functions of benefit distribution that the competitive institution serves, the concept may be easier for readers to grasp if we consider economic benefits in market competition such as market share (which is won by enterprises competing in the market). However, with benefit distribution institutions that make use of "agents" to perform such distribution, such as when the state disburses "grants" for disadvantaged communities, many people may not feel that this is a competitive institution. Actually, if the "agent" is non-corrupt, what he is doing is to distribute resources in accordance with a certain standard for benefit recipients. The value of this type of indicator fundamentally reflects the "competition" between those seeking benefits. Of course, the "agent" can also procure benefits for himself from the distribution process through methods such as bribery, but this is actually no longer a matter of benefit distribution but a problem for behavior management institutions to tackle.

(ii) The indirect function of the competitive institution is behavior promotion.

Within a competitive institution, as competitors must put in more efforts to obtain more of the object of competition, the institution also has the effect of promoting competitive (desirable) behavior. Certain managers often design competitive institutions to promote desirable behavior among target individuals. For example, some companies distribute bonuses through such an institution: the more the employee works, the more bonuses he will be eligible for. This way, employee's effort level at work can be boosted significantly.

The use of competitive institution to promote desirable behavior requires comparatively low costs.

This is because within a competitive institution the object of competition must be rare and in short supply so that individuals are willing to compete for it. Generally, costs are incurred in the promotion of desirable behavior on the part of individuals under management. However, we can set this cost (such as rewards paid out by the manager) as the object of competition and maintain its scarcity. Therefore, the quantity of the object of competition must be limited. For example, within a bonus institution, only very few individuals actually receive the awards. However, the institution can motivate all employees to work harder in order to compete for the reward.

10.1.1.3 The Measurement of Competition Determines the Institution Effects

(i) The measurement of competition refers to the indicators for comparing competitiveness of every competitor

Competitors in a competitive institution are compared using measures of competition. For example, in a competition between two persons, if the competition is about who has greater strength, the measure of competition would be "bodily strength". However, if the two are competing over who is the smarter person, the measure of competition would be "intelligence".

(ii) The measurement of competition determines the nature of competitive behavior

The measurement of competition determines the nature of competitive behaviors within a competitive institution. In some competitive institutions, if the measurement of competition gives rise to undesirable behavior, then this sort of competition often has adverse effects. Take product packaging for instance. Most times, manufacturers tend to launch their products on the market with comparatively simple packaging. At this point, the cost of manufacturing a product is comparatively low for manufacturers. However, often a manufacturer would want to increase sales of its product by putting more efforts into product packaging. As such, more investment would be made into the packaging of the product. (In the 1980s, there was a case of a manufacturer that packaged medication in carpets, effectively bribing the hospital procurement personnel in the process.) As a result, the manufacturer gained a bigger market share to the detriment of its competitors, which then led to a "packaging arms race" where manufacturers competed with each other continuously in the area. The end-result of such competition was, each manufacturer ended up with a market share close to what it had started off with, but the products have evolved from bearing fairly simple packaging to more elaborate, even "luxurious" packaging. This means the competition has added unnecessary costs to the manufacturer. Clearly, the measurement of competition for the "packaging level" has led to a form of undesirable behavior.

On the other hand, the companies have implemented a competitive institution where "only the most hardworking employee would receive a reward". Clearly, the measurement of competition used here is "the level of effort put into work", and the use of such a measurement has stimulated competition between workers in terms of "putting greater effort into work". This type of competition has given rise to individuals competing in terms of desirable behavior, which makes it healthy competition.

(iii) Conditions for good measurement of competition

Generally, the measurement of competition used is a specific characteristic of competitors but not all characteristics are suited for use as measures. When making

a selection, the institution designer should ensure that the measurement of competition meets the following conditions.

First, the measurement of competition must be based on the characteristics that competitors can change through efforts. Only this type of measurement will lead to behavioral change among individuals targeted and encourage such individuals to engage in desirable behavior.

In practice, certain competitive institutions have been effective in encouraging desirable behaviors among individuals as the measurement of competition used can be changed in outcome by the effort of competitors. Take for example some competitions linked with work performance. Since work performance can be enhanced with effort, targeted workers would then be encouraged to work harder. On the other hand, if the outcome of the measurement cannot be changed with effort, then the competition would not encourage individuals to engage in desirable behavior. For example, if we are to use factors such as the employee's age, gender, race or family background to measure competition, not only would we not be able to encourage individuals to engage in desirable behavior, we may even cause those who do not meet these requirements to give up altogether.

Second, the measurement of competition must be highly observable. As competitors seek to outdo each other in the competition, the measurement of competition must be highly observable. In other words, the measurement of competition must be obvious and overt. Otherwise, "competition outcome" would often become controversial.

Third, the competitive measurement used must encourage desirable behavior.

For individuals under management, whichever behavior that would enhance the measurement outcome would be enhanced. For the institution manager, behavior by individuals targeted for management can be divided into two categories: "desirable behavior" and "undesirable behavior". From this perspective, if a certain measurement of competition leads to desirable behavior, it is a good measurement; conversely, if it leads to undesirable behavior, it is a poor measure.

10.1.2 Sun Diagram for a Competitive Institution

The Sun Diagram for a competitive institution is shown as in Fig. 10.1.

The diagram illustrates the special situation where two individuals are competing. Here, the line that starts from e_1 points towards the positive side of the promoter (reward) r_1, which indicates that when individual 1's effort level e_1 increases, r_1 will enlarge. The arrow that departs from e_2 points towards the negative end of the promoter r_1, which means that when individual 2's effort level e_2 increases, r_1 will diminish. What the arrow indicates is that only e_2 has an impact on r_1 and that r_1 has no impact on e_2. Similarly, the line that starts from e_2 points towards the positive side of the promoter (reward) r_2, indicating that when individual 2's effort level e_2 increases, r_2 will enlarge. The line that departs from e_1 towards the negative side of the promoter r_2 means that when individual 1's effort

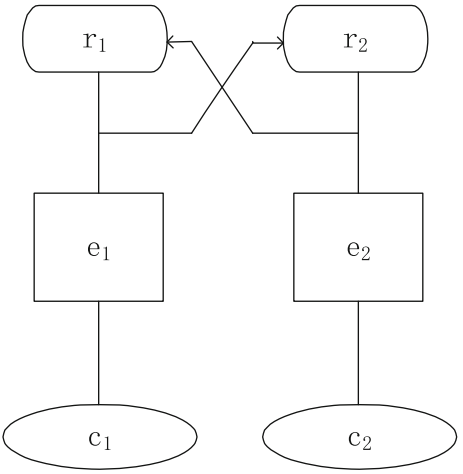

Fig. 10.1 Sun Diagram for a competitive institution (two individuals)

level e_1 increases, r_2 will diminish. The arrow indicates that e_1 has an impact on r_2 but r_2 does not have an impact on e_1. Here, we can clearly see the competitive relationship between individual 1 and individual 2.

10.1.3 Some Real-Life Examples of Competitive Institution

The competitive institution is often seen in management practice. Let us look at some real-life case studies of the competitive institution below.

Case study one: the limited-rewards institution

A company has introduced an "Outstanding Worker Award" scheme, where the top three employees by performance at the end of the year are to be awarded with a cash prize of RMB100,000 each. Everyone began to work harder after the scheme was announced. But no matter how hard everyone works, in the end only the top three employees ranked by performance would be eligible for the cash award, with the probability that each employee would get the reward increasing with their personal effort level and decreasing with the increasing effort level by other employees. Therefore, the "Outstanding Worker Award" scheme is a competitive institution and the competitive behavior in this case would be hard work.

Case study two: the project bidding institution

Some large-scale building projects are contracted out by using the bidding institution. Under this institution, the project owner releases information about the bidding through the media, with the invitation for bids accompanied by specific requirements for would-be bidders. Enterprises that meet the stipulated requirements may participate in the tender. Among participating enterprises, the enterprise

with the most reasonable quotation will win the tender. This way, for the competing enterprises, only one of them would receive the reward while the rest would receive nothing at all. And the probability that each enterprise would win the bid increases as the firm's effort level increases and decreases as the effort level of its competitor firms increases. Therefore, the construction capacity of the enterprise is a measurement of competitive behavior.

Case study three: the goods auction

To take the English-style auction for example: the agent displays the goods for auction at the behest of the seller, with interested buyers stating their offers for the goods. The offer from the would-be buyer must be higher in value than the offer from the previous bidder. This way, the price of the goods will rise continuously until no one is willing to offer an even higher price. The seller then sells the goods to the bidder with the highest price. This way, for bidders participating in the auction, only the bidder with the highest offer would be eligible for the purchase of the goods on auction. The higher the price offered by the bidder, the higher the chance that he would win the auction. However, the higher the prices offered by rival bidders, the lower the chances that he would win the bid. Therefore, bidding is a type of competitive behavior.

Case study four: the children and the lone peach

A small peach tree has just produced a single peach. A group of children all want to have the peach. In the end, the child who had first gotten to the peach was able to taste it, while the children who followed ended up with nothing. In this case, the probability of one being able to eat the peach increases the earlier one arrives and decreases the earlier the other children arrive. Therefore, arrival time is the indicator of competitive behavior here.

Case study five: product packaging

On the market, there are plenty of goods produced by different enterprises but otherwise quite similar to each other in terms of quality. Among these products, those with better packaging would have a bigger market share compared to the others. Therefore, the market share of the enterprise's product increases with the packaging costs put in by the manufacturer, and decreases with the increasing costs put in by other manufacturers. Under such circumstances, product packaging has become a competitive behavior in terms of market share.

Case study six: the arms race

Generally speaking, the stronger a nation's military might, the greater the nation's influence in the international arena. Therefore, a nation's scope of influence increases with its military investment and decreases as other nations also increase their military spending. Hence, defense spending is a competitive behavior in terms of the scope of the nation's influence.

Case study seven: standing up at the open-air cinema

In the 1970s, open-air cinemas were a common sight in Chinese villages on summer nights. Typically, attendees would come with stools and benches on which they would sit during the screening. This way, they could cool off in the open while enjoying some entertainment at the same time. However, if a particularly engaging film was being shown that day, very often one or two persons would stand up to get a better view of the screen. However, when this happened, those sitting behind would also stand up as their field of vision had been blocked. Eventually, everyone would stand up. In the end, instead of being able to watch the film sitting down everyone had to stand up in order to catch the film.

Here, the field of vision increases when one stands, and decreases when other people also choose to stand. Here, the act of standing is a competitive act.

For the first, second, third and fourth examples, the object of competition is indivisible, and the level of effort put into competitive behavior impacts the probability of one winning the object. In the fifth and sixth cases, the object of competition is divisible, and the level of effort put into competitive behavior impacts the proportion of the object of competition one would get.

10.1.4 Mathematical Models for the Competitive Institution

10.1.4.1 Object of Competition

There are two main types of objects of competition within a competitive institution.

The first type is indivisible, such as sports competition rankings and building project contracts.

The other type is divisible, such as market share or the catch in a fishing zone.

Divisible objects of competition are divided among competitors in a given ratio. Individuals of high effort levels would receive a higher proportion of the object of competition while those with lower effort levels would receive a lower proportion of the object of competition.

For objects of competition that are indivisible, typically only individuals that have won in the competition (in general, a minority) would be able to obtain the object of competition while those who have lost (typically the majority) would not be able to obtain the object. However, in order to create a standard mathematical model for the two types of objects, let us assume that the probability of obtaining the indivisible object of competition is determined by the effort level of the respective competitors. Individuals of high effort levels would have a higher chance of obtaining the object of competition while those with lower effort levels would have a lower chance of obtaining the object of competition.

This way, through the probability of individuals obtaining the object of competition, we would be able to equate the reward functions for indivisible and divisible objects of competition.

10.1.4.2 Reward Function for Competitive Behavior

To create an equilibrium-point model for the individual's behavior within a competitive institution, first assume that:

There are n competitors (individuals). There is only one object of competition of value r_0.

This way, the reward function for competitive behavior is in Eq. (10.1),

$$r_i = \frac{e_i}{\sum_{j=1}^{n} e_j} r_0 \tag{10.1}$$

While $r_0 > 0$ is the value of the item competed for. $e_i > 0$ is the effort level of individual i, and r_i is the benefit individual i can receive from engaging in competitive behavior (i.e., the expected reward from engaging in competitive behavior).

From Eq. (10.1), we can see that the maximum benefit to individual i will not exceed r_0. In other words, individual i's benefit would be $r_i \leq r_0$. When $n = 1$, that is when there is no competition, the probability of the object of competition going to individual i is 1 or 100 %. Therefore, at this time $r_i = r_0$, which is the maximum.

In Eq. (10.1), $\frac{e_i}{\sum_{j=1}^{n} e_j}$ is the probability of individual i winning the object of competition (when it is indivisible) or the proportion of the object of competition that has been won (when the object of competition is divisible).

Example 10.1 Competition through product packaging

Companies A and B produce the same type of product of equal quality and similar specifications. Both companies wish to grow their respective market share. *Ceteris paribus*, the market share of the product is determined by its packaging.

Assume that the cost of packaging is the effort level of the enterprise. Company A's packaging cost is e_1 = RMB60/Product and Company B's packaging cost is e_2 = RMB80/Product, with the object of competition here being market share. The value of the object of competition is r_0 = 100 %.

Based on Eq. (10.1), Company A's product would have a market share of:

$$r_1 = \frac{e_1}{e_1 + e_2} r_0 = \frac{60}{60 + 80} \times 100\,\% = 43\,\%$$

This means that Company A's product has a 43 % market share while Company B's product has a 57 %.

Example 10.2 Project bidding exercise

There is a tender exercise open for a building project. The winning contractor can earn RMB10 million in profits from the project, i.e., $r_0 = 1000$.

Assume that the probability that the contractor would win the tender is positively related to its competency level, that is, the more competent the contractor is at handling projects, the higher the probability that it would win the tender. However,

as the tender issuing unit may have different understanding of contractors' competency from others, the most competent contractor may not necessarily win the bidding; it simply has a higher chance compared to other competitors.

Let us assume that Contractor A and Contractor B are participating in the tender exercise. The competency level of the contractor is expressed in numerical form, with the competency value of Contractor A at 5 and the competency value of Contractor B at 4. Based on Eq. (10.1), Contractor A's expected reward from participating in the tender exercise is:

$$r_1 = \frac{e_1}{e_1 + e_2} r_0 = \frac{5}{5+4} \times 1000 = 0.56 \times 1000 = \text{RMB5.6 million}$$

In this example, the probability of Contractor A winning the bid is 56 %.

10.1.4.3 Cost of Competitive Behavior

The individual i's cost of competitive behavior c_i is a function of level of effort put into competitive behavior e_i, with the mathematical model as follows:

$$c_i = e_i \tag{10.2}$$

Equation (10.2) means that the effort levels of competitors can be expressed in the form of the costs incurred in the contractors' participation in the tender exercise. For example, when the manufacturer engages in competition for greater market share for its product, the cost of advertising, promotions, packaging, etc., reflects its effort level in competitive behavior. The higher these costs, the higher the effort level on the part of the manufacturer. To take another example, in the competition for "greater field of vision" in an open-air cinema, the field of vision for individuals sitting down leisurely would be smaller. We can regard such behavior as bearing comparatively lower effort cost. While those who have opted to watch the movie standing up would have expended more energy, they would also enjoy a greater field of vision. Here, their competitive cost or effort level in competition is comparatively higher.

10.1.4.4 The Behavior Utility Function

When n individuals are competing with each other, set the function of individual i's utility u_i as Eq. (10.3).

$$u_i = r_i - c_i = \frac{e_i}{\sum_{j=1}^{n} e_j} r_0 - e_i$$

i.e.:

$$u_i = \frac{e_i}{\sum_{j=1}^{n} e_j} r_0 - e_i \tag{10.3}$$

For Eq. (10.3), we should note that when calculating utility we must standardize the units for reward r_i and cost c_i. Typically, they are standardized to economic value. For instance, in example 10.1, the reward unit is market share while the cost is the economic cost of packaging. If we are to use Eq. (10.3) to calculate utility, we would need to convert the reward unit from market share to the corresponding profits the enterprise would stand to make with such a market share. This way, both reward and cost are expressed in economic terms and we can then proceed to calculate the corresponding utility.

10.1.5 Equilibrium for the Level of Effort Put in by Individuals Within the Competitive Institution

10.1.5.1 The Equilibrium for Effort Level in Competitive Behavior

Seek the partial derivative of the effort level e_i with the function of utility equation (10.3), and take the partial derivative 0 as follows:

$$\frac{\partial r_i}{\partial e_i} = \frac{\sum_{j=1}^{n} e_j - e_i}{\left(\sum_{j=1}^{n} e_j\right)^2} r_0 - 1 = 0, \text{ i.e., } r_0\left(\sum_{j=1}^{n} e_j - e_i\right) = \left(\sum_{j=1}^{n} e_j\right)^2.$$

Of which, $\left(\sum_{j=1}^{n} e_j\right)^2 = \left(e_i + \sum_{j\neq i} e_j\right)^2$, $r_0\left(\sum_{j=1}^{n} e_j - e_i\right) = r_0 \sum_{j\neq i} e_j$. Therefore, we have $r_0 \sum_{j\neq i} e_j = \left(e_i + \sum_{j\neq i} e_j\right)^2$.

We then derive the equilibrium point of individual i's effort level e_i:

$$e_i^* = \sqrt{r_0 \sum_{j\neq i} e_j} - \sum_{j\neq i} e_j \tag{10.4}$$

10.1.5.2 Factors that Impact the Equilibrium Level in Competitive Behavior

Let us now look at the factors that impact the equilibrium point e_i^* of individual i's effort level.

(i) The higher the value of the object of competition, the higher the equilibrium level

From Eq. (10.4), we can see there is a positive linear relationship between e_i^* and $\sqrt{r_0}$. This means that the higher the value of the object of competition r_0, the higher the equilibrium level of effort e_i^* by individual i.

(ii) Under specific circumstances, the higher the effort level of other players, the higher the level of effort put into competitive behavior by the individual

We now look at the effort level of other individuals apart from i as well as the impact of $\sum_{j\neq i} e_j$ on the effort equilibrium point e_i^* for individual i. Seek the partial derivative of $\sum_{j\neq i} e_j$ with the function of utility equation (10.4) as follows:

$$\frac{\partial e_i^*}{\partial(\sum_{j\neq i} e_j)} = \sqrt{\frac{r_0}{\sum_{j\neq i} e_j}} - 1 \tag{10.5}$$

When $r_0 > \sum_{j\neq i} e_j$, $\frac{\partial e_i^*}{\partial(\sum_{j\neq i} e_j)} > 0$. It means that the effort equilibrium level e_i^* for individual i will increase as the sum total of effort by other individuals $\sum_{j\neq i} e_j$ increases. This indicates the competitiveness of effort level. Considering Eq. (10.2), i.e., when the cost of competitive behavior for each individual equals the individual's effort level in competition, $r_0 > \sum_{j\neq i} e_j$ means that the effort level on the part of other individuals apart from individual i is not yet excessive and that the total cost of their competitive behavior is lower than the value of the object of competition. At this point, if individual i were to join the competition, the value of the portion of the object of competition he would receive would be greater than his cost, which means that his competitive behavior would bring positive utility.

When $r_0 \leq \sum_{j\neq i} e_j$, $\frac{\partial e_i^*}{\partial(\sum_{j\neq i} e_j)} \leq 0$. It means that the equilibrium point e_i^* for the effort level of individual i would not increase in tandem with the sum of the effort levels of other competitors $\sum_{j\neq i} e_j$. Considering Eq. (10.2), i.e., when the cost of competitive behavior for each individual equals their effort level in competition, $r_0 \leq \sum_{j\neq i} e_j$ means that the effort level on the part of other individuals apart from individual i has become excessive and that the total cost of their competitive behavior has become greater than the value of the object of competition. At this point, if individual i were to join the competition, the value of the portion of the object of competition he would receive would no longer be greater than his cost, which means that his competitive behavior would not bring positive utility.

10.1.5.3 The Greater the Number of Competitors, the Lower the Equilibrium Point for Effort Level

The impact of the number of competitors on effort equilibrium is a key issue as it is significant in terms of how we can tackle unhealthy competition. This issue will be examined in the following section.

10.2 The Impact and Significance of the Number of Individuals in Symmetric Competition on the Equilibrium Point

10.2.1 Symmetric Competition

Let us now look at the scenario of symmetric competition. Symmetric competition happens when all competitors are of the same nature. Under such circumstances, when competition hits equilibrium, each individual would have put in an equal amount of effort. This is also what we usually call an "evenly-matched" scenario.

The concept of symmetric competition has tremendous research significance. This is because in many cases competitors within an institution are quite similar to each other. If there is a vast difference between individuals, very often there would not even be a competition at all. Take for example competition between countries, between enterprises, between schools, between regions, between employees, etc.

10.2.2 Competition Is Most Intense Where There Are Only Two Competitors in the Field

Assuming symmetric competition, re-write effort equilibrium Eq. (10.4) into Eq. (10.6).

$$\begin{aligned} e_i^* &= \sqrt{r_0 \sum_{j \neq i} e_j} - \sum_{j \neq i} e_j = \sqrt{r_0 (n-1) e_i^*} - (n-1) e_i^* \\ n e_i^* &= \sqrt{r_0 (n-1) e_i^*} \\ n^2 (e_i^*)^2 &= r_0 (n-1) e_i^* \\ e_i^* &= \frac{(n-1)}{n^2} r_0 \end{aligned} \tag{10.6}$$

Based on Eq. (10.6), when $n \to \infty$, $e_i^* \to 0$. We can hence see that the greater the number of competitors, the lower the equilibrium point for effort level. This result can help to explain a few phenomena. Take for instance an office with five or six employees. When it is necessary to select one of them as leader, the ensuing competition would tend to be intense. However, if we are to run the same competition on a factory floor with five or six hundred workers, we would see that competition in this group would not be intense at all.

Some specific outcomes from Eq. (10.6) are very interesting and aligned with what goes on in management practice.

For example, when there is only one player in the field, i.e., $n = 1$, clearly, the object of competition would be gained without any need for competition. Based on

Eq. (10.6), $e_i^* = \frac{(1-1)}{1^2} r_0 = 0$, i.e., when the effort level in competition is 0, no competitive behavior would be needed either.

When there are two individuals, $n = 2$. $e_i^* = \frac{(2-1)}{2^2} r_0 = \frac{1}{4} r_0$.

When there are three individuals, $n = 3$. $e_i^* = \frac{(3-1)}{3^2} r_0 = \frac{2}{9} r_0$.

When there are four individuals, $n = 4$. $e_i^* = \frac{(4-1)}{4^2} r_0 = \frac{3}{16} r_0$.

… when $n \to \infty$, $e_i^* \to 0$.

It is clear that as the number of competitors in the field increases, the effort equilibrium level for each competitor becomes increasingly lower at a rather significant rate.

What is interesting about Eq. (10.6) is that when $n = 1$, we get $e_i^* = 0$ while when $n = 2$, we get $e_i^* = \frac{1}{4} r_0$. This is when the equilibrium point for effort level in competitive behavior is the highest. And as n increases, the equilibrium point for effort level in competitive level e_i^* becomes lower and lower.

That is to say, in a situation where the value r_0 of the object of competition (i.e., the impact of a change in the number of competitors n on competition intensity by assuming the object of competition remains unchanged), the competition is most relaxed when there is only one player in the field. At this point, the level of effort put into competition is 0. Competition is the most intense when there are only two competitors in the game. As for groups where the total number of competitors is greater than two, the intensity of competition would be between that of a group of $n = 1$ and of a group of $n = 2$, and will decrease as the total number of competitors increases. When $n \to \infty$, $e_i^* \to 0$. This would be akin to a situation where a prime minister is to be selected from a population of one billion people. It would be an affair that no one would be fussed about and there would be very little effort in competition.

During the Cold War, as there were only two superpowers in the world at that time, the U.S.A. and the Soviet Union, the rivalry between the two became extremely intense. Indeed, the two had brought the world to the brink of nuclear war in the 1950s. This is the intense rivalry caused by $n = 2$. Later on, with the break-up of the Soviet Union, a multi-polar international landscape began to emerge and international rivalries showed signs of abating.

In the same vein, based in Eq. (10.6), competition in a perfectly competitive market as described in theories of economics would not be intense as in an oligopolistic market as there is a large number of enterprises of various sizes in this market! Experience also tells us that in a perfectly competitive market with competitors of all sizes, behavior triggered by competition are all to do with the enhancement of product quality or the lowering of costs. These are normal behaviors. However, in an oligopolistic market, the intense rivalry between a small number of competitors would often lead to unhealthy competitive behavior such as slander and cutthroat price competition.

Competition is most intense when there are only two competitors in the field. The greater the number of competitors, the less intense competition becomes. This conclusion runs contrary to the wisdom of classic economics and management

theory that the greater the number of competitors in the field, the more intense the competition. As this conclusion is derived from a stringent mathematical model, it is a key scientific finding of this book.

10.2.3 The Greater the Number of Competitors, the Smaller the Benefits

For a situation of symmetric competition, substitute the equilibrium equation (10.6) into the utility equation for competitive behavior (10.3). We then have:

$$u_i = \frac{e_i}{\sum_{j=1}^{n} e_j} r_0 - e_i = \frac{\frac{(n-1)}{n^2} r_0}{n\frac{(n-1)}{n^2} r_0} r_0 - \frac{(n-1)}{n^2} r_0 = \frac{r_0}{n} - \frac{(n-1)}{n^2} r_0$$
$$= \frac{nr_0 - (nr_0 - r_0)}{n^2} = \frac{r_0}{n^2}$$

Therefore, the utility for various competitors participating in symmetric competition is:

$$u_i = \frac{r_0}{n^2} \tag{10.7}$$

Equation (10.7) shows that the greater the number of competitors, the lower the utility gained from competitive behavior for each individual. Therefore, Eq. (10.7) provides a good explanation for why in industries where the entry threshold is low (i.e., industries where there are many market players) operating profits tend to be very low.

10.2.4 The Practical Significance of the Symmetric Competition Equation in Terms of Employee Motivation

The use of the competition mechanism to encourage employees to work harder is a very cost-effective approach. Hence, in practice we often see competition used as a way of motivating employees. This institution can motivate all employees to work harder for the chance to get a reward while only a small minority, the most hard-working employees, would need to be rewarded. In some government agencies, very few individuals receive job promotions for outstanding work performance. However, as job promotions are objects of competition, they can motivate employees to work harder in order to earn these promotions.

However, the equation for effort equilibrium in competition (10.6) suggests the potential issues that may arise when using a competitive institution for employee motivation. One issue that may impact motivational effect would be the number of competitors in the group: if n is too large, the effect would be compromised. Imagine if 500 persons were to compete for a single reward. The majority of these 500 people would give up trying altogether.

The solution to this problem would be to provide tiered rewards. For instance, groups of five persons each could be formed, with a consolation prize-type reward made available. Then, a third prize could be given out for every five employees who have been qualified for the consolation prize, a second prize given out for every five persons who have been qualified for a second prize, and a first prize given out for every five persons who have been qualified for a second prize. This way, while there would be 125 individuals vying for the first prize, as the rewards are tiered with a consolation prize given to one out of every five persons, this would lead to the situation where $n = 5$. Therefore, n would not appear to be very great and all workers could still be motivated effectively.

10.2.5 The Significance of the Equation of the Symmetric-Competition Equilibrium Point for Tackling Unhealthy Competition

In management practice, we often see unhealthy behavior that can be hard to tackle.

For example, the product packaging problem is a kind of undesirable and unhealthy competition. In order to win greater market share, product manufacturers have incurred ever higher costs in over-packaging their products, leading to higher product prices and material wastage. The resulting garbage has also exacerbated the environmental pollution. For a long time, among other measures the government has chosen to deal with the over-packaging issue through the enactment of laws that penalize such behavior as well as encouraging consumers to contact the authorities with tip-offs of such behavior. However, over the years, the problem of over-packaging remains and appears to be a stubborn beast. Clearly, the root of this situation lies in the intense competition between manufacturers for market share.

Based on Eq. (10.6), the greater the number n of competing manufacturers, the lower the level of effort put into competition. Therefore, one possible measure would be to increase the number of manufacturers in the field to lower effort level in order to eliminate the problem of over-packaging. Therefore, Eq. (10.6) has provided us with another way of thinking about the matter: if we detect the problem of product over-packaging in a certain industry, we may then work to eliminate any oligopoly present in the market and introduce a significant number of small and medium enterprises into the competition. This way, the market share gained by the manufacturer through more lavish packaging would shrink, which means that the profits of doing so would decrease. As such, the manufacturer would no longer have

an interest in over-investing in product packaging. The conclusion is also useful for tackling other forms of unhealthy competition (such as price-slashing) between enterprises. The introduction of a large number of competitors into the competition will reduce the reward from engaging in adverse competition and hence reduce the interest of enterprises in engaging in such behavior.

10.3 Synchronized Competition Model and Unnecessary Costs

10.3.1 The Equilibrium Point in Synchronized Competition

Synchronized competition happens when the effort level of all individuals in competition is required to change in a synchronized manner during competition.

For synchronized competition to happen, there must be a third-party authority present, and this authority must be powerful enough to fully control the effort levels of n individuals so that they would engage in competition at a synchronized effort level. In other words, at any moment during the competition process all competitors must put in the same level of effort in competitive behavior: $e_1 = e_2 = \cdots = e_i = \cdots = e_n$. At this point, the utility function of individual i can be re-written from (10.3) to (10.8).

$$u_i = \frac{e_i}{\sum_{j=1}^{n} e_j} r_0 - e_i = \frac{e_i}{ne_i} r_0 - e_i = \frac{r_0}{n} - e_i$$

i.e.:

$$u_i = \frac{r_0}{n} - e_i \tag{10.8}$$

Clearly, with the rule that $0 \leq e_i$, when $e_i = 0$, the value u_i is maximized. The effort equilibrium equation for synchronized competition is (10.9) while the utility equation for competitors is (10.10).

$$e_i^* = 0 \tag{10.9}$$

$$u_i = \frac{r_0}{n} \tag{10.10}$$

That is to say, during synchronized competition, the object of competition is distributed in $\frac{1}{n}$ shares (when the object is divisible) or at the probability of $\frac{1}{n}$ (when the object is not divisible) to each individual. Therefore, it actually has nothing to do with everyone's effort level during competition. Under such circumstances, the optimal effort level for individual utility maximization is of course 0, i.e., with no

competitive behavior undertaken as such behavior incur behavioral cost while not bringing in additional rewards.

10.3.2 Unnecessary Costs in Competition

Let us now compare the utility difference between symmetric competition and synchronized competition.

Compare the individual utility equations (10.7) and (10.10) for the equilibrium points in effort levels in a scenario with two individuals, as

$$\frac{r_0}{n} > \frac{r_0}{n^2}, \quad (n > 1)$$

Thus, in a situation where there is synchronized competition the utility of the individual is higher than individual utility in a situation of symmetric competition.

The cause of this difference lies mainly in the different ways in which symmetric competition and synchronized competition come about.

In symmetric competition, each participating individual is similar to its competitors in terms of competitiveness. However, the effort level shown by each competitor during competition would differ. However, as all competitors are of a similar competitive level, eventually at equilibrium the effort level shown by each individual in competitive behavior would definitely be the same.

In the situation of synchronized competition, during the process of competition effort level is the same across all competitors. This requirement for "similar levels of effort" is even stronger than in the case of symmetric competition. In symmetric competition, only the final effort equilibrium level for all competitors would be the same, while in synchronized competition the effort level for all players must be similar at all times. Of course, this would also entail the final effort equilibrium level for all competitors being the same.

The "similar levels of effort" in the process of synchronized competition means that the competitive behavior of all players is controllable and also implies the existence of an impartial authority. This authority can achieve the following two objectives through the limitation of effort level on the part of all players:

First, the authority can equalize the utility for all competitors (when all competitors are equally competitive). Second, the authority can maximize utility for all competitors.

On the other hand, we see from Eq. (10.9) that during synchronized competition, $e_i^* = 0$ is where individual utility is maximized. This means that if competitive behavior is controllable to the extent of ensuring synchronized competition, then from the perspective of maximizing competitor utility competitive behavior should be eliminated completely!

The significance of this conclusion is that when competition arises between multiple individuals of a similar nature (i.e., these competitors are evenly matched),

very often such competition would only lead to everyone incurring the cost of competition $c_i = e_i$ in vain and making competition necessary for obtaining of the object of competition when no competition was required in the first place. This is just like the case of the open-air cinema: the audience had to resort to "standing up to better see the film", turning an experience where sitting would suffice to one where standing up was necessary in order to continue watching the film. Other forms of competition, such as over-packaging and arms race, also prove this principle. And this is the unnecessary cost in unhealthy competition. The term "pyrrhic victory" describes precisely this kind of situation.

10.4 Inhibiting Unhealthy Competition: A Competitive Institution with a "cop"

There are two key causes of the competitive institution: first, it has been designed to enhance desirable behavior among targets of management, and second, it occurs naturally due to various objective factors.

In naturally-occurring competitions, some are of an adverse or unhealthy nature. Such competition reinforces undesirable behavior and poses a certain risk. As such, societies have tried to come up with institutions to tackle undesirable behaviors, and a classic form of such institutions is the competitive institution "with cop".

The "cop" we are referring to here is not a vocation but an abstract role played by a third-party who meets the following three conditions:

1. He is a third-party who is not engaged in the competition, and has no relationship of interest with any competitor;
2. He regards the total benefits earned by all competitors as benefits to himself; and
3. He has the ability to penalize unhealthy or undesirable competitive behavior on the part of competitors.

10.4.1 The Concept and Sun Diagram of the Competitive Institution "with cop"

10.4.1.1 The Concept of the Competitive Institution "with cop"

The competitive institution "with cop" is an institution designed to lower the level of effort put into undesirable competitive behavior.

In management practice, as adverse competition is detrimental to group interests, very often there are agencies or professions that fulfill the role of the "cop". Generally, they allow mild competition but forbid excessive competition in order to avoid the damage that adverse competition can bring. For instance, when the

audience is watching a film inside a movie theater, venue personnel would oftentimes also act as "cops" that maintain venue order. Members of the audience may crane their necks to see the screen better, but may not stand up and thus block the view of others. The commerce and industry agencies also act as "cops" in the management of competition between enterprises. These agencies permit enterprises to moderately package products but do not permit the practice of over-packaging. The civil police in the public security agencies act as the "true cops" that manage property competition, allowing residents to lock their doors and windows in order to deter burglars but not to possess firearms. Effective management by these cops has led to the long-term stability of society without sacrificing competition altogether.

With regard to the Sun Diagram, the key feature of the competitive institution "with cop" is the addition of a suppressor equipped with an observer into the institution. This additional component is the "cop". This response characteristic of this suppressor vis-à-vis competition effort level is, when the effort level exceeds a certain point a negative reward would be generated for the actor. This negative reward is also exacerbated as the effort level continues to rise. This way, the individual's competitive behavior would be suppressed and limited to within acceptable levels.

10.4.1.2 Sun Diagram of the Competitive Institution "with cop"

The Sun Diagram for a competitive institution "with cop" is shown as in Fig. 10.2.

The diagram illustrates a special situation where two individuals are competing. Here, the line that departs from e_1 points towards the observer p_1 of the "cop", indicating that when individual 1's effort level e_1 exceeds the maximum level permitted by the "cop", the "cop" has a probability of p_{11} of discovering the

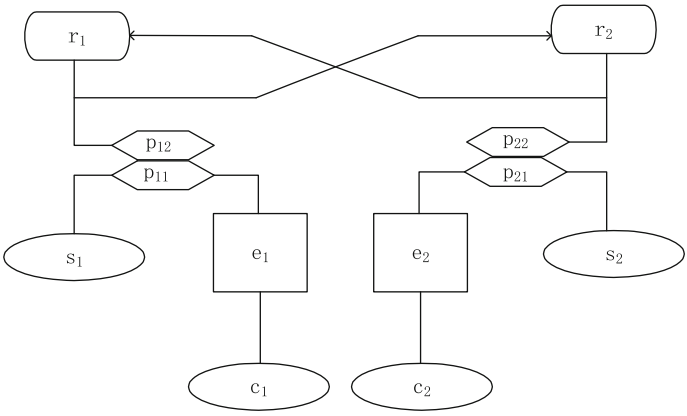

Fig. 10.2 Sun Diagram for competitive institution "with cop" (two individuals)

individual's undesirable behavior. If discovered, the individual would receive a negative reward s_1 meant to suppress such behavior. In the case of the movie theater, it would mean that the offending member of the audience would be driven out, in the case of product over-packaging, the offending enterprise would be fined, etc. At the same time, adverse competition has a probability of $p_{12} = 1 - p_{11}$ of "not being discovered by the cop". At this point, the reward from unhealthy competition would be r_1 and at the same time the reward r_2 to e_2 would decrease. Similarly, the line that departs from e_2 points towards the observer p_2 of the "cop", indicating that when individual 2's effort level e_2 exceeds the maximum level permitted by the "cop", the "cop" would have a probability of p_{21} of discovering such behavior. If discovered, the individual would receive a negative reward s_2 designed to suppress such behavior. At the same time, adverse competition has a probability of $p_{22} = 1 - p_{21}$ of "not being discovered by the cop". At this point, the reward from unhealthy competition would be r_2 and at the same time the reward r_1 to e_1 would decrease.

10.4.2 Equilibrium Model and the Equilibrium Point of Effort Level

(i) The "cop function"

The "cop function" expresses the degree to which individual i would be penalized when he engages in undesirable behavior and is detected by the "cop". In this book, we assume that the cop function is a piecewise linear function of competitor effort level, i.e., Eq. (10.11):

$$s_i = \begin{cases} ke_i, & e_i \geq l \\ 0, & e_i < l \end{cases} \tag{10.11}$$

In the equation, $s_i > 0$ is the penalty the cop metes out to individual i, with $e_i > 0$ individual i's effort level in competitive behavior and $l > 0$ the maximum effort level permitted by the cop. $k > 0$ is the effort level penalty rate as meted out by the cop.

(ii) Reward function

The reward function is the positive reward received by individual i through undesirable behavior when the undesirable behavior by the individual has not been penalized by the "cop". In this book, we assume that the reward function still takes the form of Eq. (10.1):

$$r_i = \frac{e_i}{\sum_{j=1}^{n} e_j} r_0$$

(iii) The cost of competitive behavior

For the cost of competitive behavior in a competitive institution "with cop", we continue to use Eq. (10.2), i.e., the cost of competitive behavior by individual i is:

$$c_i = e_i$$

(iv) The utility function in a competitive institution "with cop"

There are two types of situations for the utility function for individual i in a competitive institution "with cop".

When $e_i < l$, the utility function is the same as Eq. (10.3), i.e.:

$$u_i = \frac{e_i}{\sum_{j=1}^{n} e_j} r_0 - e_i$$

When $e_i \geq l$, the utility function of competitive behavior is expressed in Eq. (10.12). The reader should note that the probability of undesirable behavior being detected by the "cop" is p_{i1} and the probability of not getting caught is $(1 - p_{i1})$:

$$\begin{aligned} u_i &= (1 - p_{i1}) r_i - p_{i1} s_i - c_i = (1 - p_{i1}) \frac{e_i}{\sum_{j=1}^{n} e_j} r_0 - p_{i1} k e_i - e_i \\ &= (1 - p_{i1}) \frac{e_i}{\sum_{j=1}^{n} e_j} r_0 - (p_{i1} k + 1) e_i \end{aligned}$$

i.e.:

$$u_i = (1 - p_{i1}) \frac{e_i}{\sum_{j=1}^{n} e_j} r_0 - (p_{i1} k + 1) e_i \tag{10.12}$$

(v) The equilibrium level for effort level in competitive behavior

Seek the partial derivative of the effort level e_{i} with the function of utility Eq. (10.12), and take the partial derivative 0 as follows:

$$\begin{aligned} &\frac{\partial r_i}{\partial e_i} = (1 - p_{i1}) \frac{\sum_{j=1}^{n} e_j - e_i}{\left(\sum_{j=1}^{n} e_j\right)^2} r_0 - (p_{i1} k + 1) = 0, \quad \text{i.e.,} \\ &r_0 (1 - p_{i1}) \left(\sum_{j=1}^{n} e_j - e_i\right) = (p_{i1} k + 1) \left(\sum_{j=1}^{n} e_j\right)^2 \end{aligned}$$

Of which, $\left(\sum_{j=1}^{n} e_j\right)^2 = \left(e_i + \sum_{j \neq i} e_j\right)^2$, $r_0 \left(\sum_{j=1}^{n} e_j - e_i\right) = r_0 \sum_{j \neq i} e_j$. Therefore, we have:

$r_0(1-p_{i1})\sum_{j\neq i}e_j=(p_{i1}k+1)(e_i+\sum_{j\neq i}e_j)^2$, to derive individual i's effort equilibrium point:

$$e_i^*=\sqrt{\frac{r_0(1-p_{i1})\sum_{j\neq i}e_j}{(p_{i1}k+1)}}-\sum_{j\neq i}e_j \tag{10.13}$$

If the individual is rational, then his effort level will find an equilibrium point determined in Eq. (10.13).

Comparing Eqs. (10.13) and (10.4) (i.e., $e_i^*=\sqrt{r_0\sum_{j\neq i}e_j}-\sum_{j\neq i}e_j$), as $1\geq p_{i1}\geq 0$, $1\geq 1-p_{i1}\geq 0$, $r_0(1-p_{i1})\sum_{j\neq i}e_j\leq r_0\sum_{j\neq i}e_j$; $k>0$, $p_{i1}k\geq 0$, $p_{i1}k+1\geq 1$, hence:

$$\sqrt{\frac{r_0(1-p_{i1})\sum_{j\neq i}e_j}{(p_{i1}k+1)}}-\sum_{j\neq i}e_j\leq\sqrt{r_0\sum_{j\neq i}e_j}-\sum_{j\neq i}e_j \tag{10.14}$$

The individual effort level of competitive behavior in a competitive institution "with cop" is lower than the individual effort level of competitive behavior in a competitive institution "sans cop".

Examining Eq. (10.14), we discover that only when $p_{i1}=0$, or when the cop's observation of unhealthy behavior is 0, the effort level in the competitive institution "with cop" would be the same as the effort level in the competitive institution "sans cop". Actually, if $p_{i1}=0$, it means that the cop is not penalizing any sort of unhealthy competition and therefore serves zero purpose. Under this circumstance, the competitive institution "with cop" has deteriorated to become a competitive institution "sans cop", which is why the equilibrium in both scenarios would be the same.

10.4.3 Key Conclusion: The Social Significance of the "cop"

From the above analysis, we see that the significance of the cop is not simply to "maintain law and order" as we commonly understand but also to inhibit unhealthy competition between individuals so that the equilibrium points of them would not deviate too far from the Pareto Optimality. In practice, competitive institutions "with cop" are commonly deployed in another indication of the significance of this type of institution.