Sketch of Reflexive Game Theory

Vladimir A. Lefebvre
School of Social Sciences
University of California, Irvine 92697

Abstract

This work gives a short description of the theory of agents endowed with reflexion and some applications of this theory for modeling human choice. Unlike classical game theory, this theory has been constructed for prediction of the agents' real choices rather than the optimal decision making.

Contents

1 **Reflexive Structure** 2

2 **Reflexive Structure as a Computational Scheme** 3

3 **Representing Reflexive Structures as Exponential Formulae** 4

4 **Vertical and Horizontal Operations** 4

5 **The Subject, Free Will, and Freedom of Choice** 5

6 **Relation between the Realist's Equation and the Logistic Differential Equation** 8

7 **The Subject's Reflexive Structure** 9

8 **Two Aspects of Choice** 10

9 **Subject with a Model of the Situation** 11

10 **Examples of Modeling the Process of Choice** 16

11 **Surprise Actions** 19

12 **Ideal Behavior** 20

13 **Imitation of the Other** 21

14 **Subject with a Quadratic Model of the Situation** 23

15 **Streams of Consciousness and Difference Equations** 25

16 **Streams of Consciousness and Acts of Awareness** 26

17 **The Moment of Choice** 29

18 **Completion of the Classical Theory of Zero-Sum 2x2 Games** 30

19 **Reflexive Games with Saddle Point** 32

20 **Risk and Caution** 33

21 **Generalization to Non-Zero-Sum Games** 35

22 **Model of the Reflexive Subject and Experimental Psychology** 37

23 **References** 41

Introduction

agents endowed with reflexion; expanding the idea of reflexion: a graphic metaphor; reality-for-the-agent

Reflexion is traditionally understood as the human ability to take the position of an observer in relation to one's own thoughts and feelings. We will expand this concept to include the ability to mentally reconstruct another person's thoughts and feelings (Lefebvre, 1965a; 1967, 1972). Our main idea is

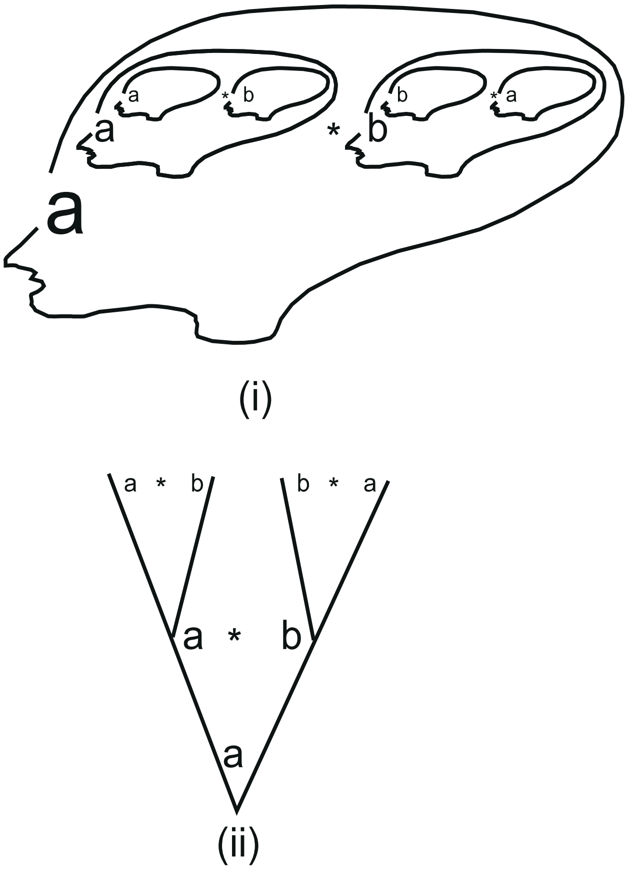

Figure 1: A reflexive agent with images of the self and of the other; each image also contains images of the self and of the other.

illustrated by the cartoon and a tree diagram in Fig.1.

The big face represents a reflexive agent. The two faces inserted into the big face represent an "image of the self" and an "image of another agent." The lower symbol * denotes an image of the agents' relations. Their status of "images" is only for us, external observers; for the agent, they play the role of reality; they are the "real me," the "real other," and the "real relation." The four small faces play the role of images of reality, which, from the agent's point of view, belong to him (left pair of faces) and to the other agent (right pair of faces).

The tree structure (Fig.1ii) depicts the agent's reflexion. The lowest knot, the tree root, corresponds to the agents' observable behavior. Two knots on the second tier with an asterisk between them symbolize reality-for-the-agent. The knots and asterisks on the third tier symbolize representations of reality by the first and second agents, from the first agent's point of view.

The agent portrayed in Fig.1i not only knows that he is in a relation with the second agent; he also knows that he knows this and that the other knows it. We can say that this agent is *aware* of the situation (Lefebvre, 1965b; 1966).

1 Reflexive Structure

a tree of images; relations between images; hierarchy of realities

The idea described in Introduction can be generalized to any finite, and even infinite, number of tiers (Fig.2).

Every node is an observer. The root of the tree corresponds to the agent and other nodes to the images and images of images.

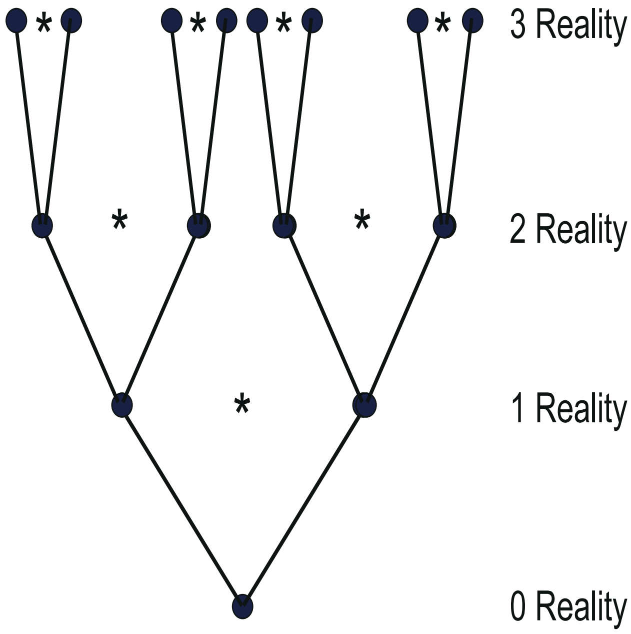

Figure 2: Reflexive structure.

The reality-for-us is the root of this tree. All other nodes have the status of representing reality-for-us. The reality-for-the-agent is the observers on the second tier and relation * between them. The reality for those observers consists of the observers on the third tier and the relations * between them. And, in general, the reality for the observers on n-th tier consists of the observers on $(n+1)$th tier and the relations $*$ between them. If a branch ends at a particular observer, this means that he does not have images and is represented in the structure only by his behavior (Lefebvre, 1972; 1980; 1992).

2 Reflexive Structure as a Computational Scheme

a tree-scheme of computation: horizontal and vertical operations; types of computations

Our next idea consists of using the reflexive structure (see Fig.2) as a tree-scheme of computations describing the real processes of cognitive computations which leads to human actions (Lefebvre et al., 1969; Lefebvre, 1982).

We can illustrate the principle of tree-like computation by the following arithmetic example:

$$2^{2^{1+1}+3^{1+0}} \to 2^{2^2+3^1} \to 2^{4+3} \to 2^7 \to 128. \quad (2.1)$$

The essence of tree-like computation consists of dividing all the computational operations into two classes: *horizontal* and *vertical*. The horizontal operation is performed first, then the vertical (see Fig.3):

In order to link the reflexive structure with computational processes we have to put every triangle of the tree shown in Fig.2 into correspondence with a pair of operations, one horizontal and one vertical. Next we formulate a rule linking every node with the element from the set on which computation is being performed. Then, beginning from the upper ends, operations are sequentially carried out and result in a final element. Infinite reflexive structures have meaning only in cases where a procedure of limit transformation is defined.

In the reflexive models which have been constructed until now, the following sets were used:

- all real numbers
- real numbers from the interval [0,1]
- Boolean elements 1 and 0
- Logical judgments

The set of all real numbers was used to model the cognitive process of generating the inner values (Lefebvre et al., 1969); real numbers from the interval [0,1] were used for modeling probabilistic choice and categorization (Lefebvre, 1982; 1992b; Schreider, 1994; Krylov, 1994; Miller, 1997a,b,c); Boolean elements 1 and 0 were used in modeling moral choice and the representation of ethical systems (Lefebvre, 1977a; 1979; 1980); logical judgments were used for modeling the processes of critical reflexion and decision making in different ethical systems (Lefebvre, 1984; Taran, 1997; Schreider, 1998a).

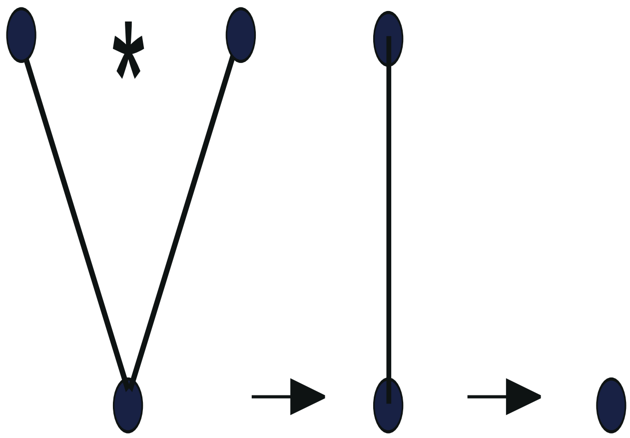

Figure 3: An elementary tree-like computation. Nodes correspond to the elements of the set to which computational operations are applied. The symbol * corresponds to the horizontal operation, and the vertical line to the vertical operation.

3 Representing Reflexive Structures as Exponential Formulae

constructing formulae; an example

It is helpful to represent reflexive computations by exponential formulas, because their structure is similar to arithmetic operations and exponent (Lefebvre, 1977c; 1984). A formal scheme for constructing such formulas is given below.

1. Initial elements:

a' letter a with an apostrophe,

b' letter b with an apostrophe,

$*$ asterisk.

Two letters combined by the symbol $*$ are called a *pair*.

2. Operations with the initial elements.

We can replace each apostrophe by a letter with an apostrophe, by a pair, or leave it as is.

3. Construction of formulas:

(i) a letter with an apostrophe or a pair is written,

(ii) the operation of substitution described in point 2 is repeated a finite number of times.

The result we obtain is a graphical object called a *formula*. If the process begins with the element a', the formula is called *agent* A; if it begins with b', it is called *agent* B, and if it begins with a pair, it is called *situation*. Let us agree to omit the apostrophe, but to keep it in mind. In addition, each letter and each asterisk can be assigned an individual number. The process of constructing a formula corresponding to Fig.1 is performed as follows:

(1) Choose element a'

(2) Substitute pair $a' * b'$ for an apostrophe:

$$\begin{array}{l} \quad a' * b' \\ a \end{array} \tag{3.1}$$

(3) Substitute pair $a' * b'$ for the left apostrophe and $b' * a'$ for the right one:

$$\begin{array}{l} \qquad a' * b' \qquad b' * a' \\ \quad a \qquad\quad * \, b \\ a \end{array} \tag{3.2}$$

(4) Eliminate the apostrophes and assign numbers to the letters and symbols:

$$\begin{array}{l} \qquad a_3 *_3 b_3 \qquad b_4 *_4 a_4 \\ \quad a_2 \qquad\quad *_2 \, b_2 \\ a_1 \end{array} \tag{3.3}$$

4 Vertical and Horizontal Operations

Boolean function as operation; gamma-algebra; exponential form

From a purely mathematical point of view, vertical and horizontal operations can be represented by any functions. Their selection can be based on a special psychological or anthropological study. In one of the first particular models describing moral choice, the vertical operation was a Boolean function called implication:

$$f(a, b) = a \vee \overline{b}, \tag{4.1}$$

or, in exponential form:

$$f(a, b) = a^b. \tag{4.2}$$

The horizontal operations used were disjunction and conjunction:

$$f_1(a, b) = a \vee b, \tag{4.3}$$

$$f_2(a, b) = a \& b. \tag{4.4}$$

The result we obtained was a Boolean model of ethical cognition, which we used to suggest the hypothesis that only two different ethical systems can exist (Lefebvre, 1977c; 1980).

The next step consisted of replacing Boolean algebra with a special algebraic object called the gamma-algebra (Lefebvre, 1982; 1992). This object can be modeled with the numbers from the interval $[0, 1]$, for which one unary and three binary operations are defined:

$$\overline{x} = 1 - x, \tag{4.5}$$

$$F(x, y) = 1 - y + xy, \tag{4.6}$$

$$x \oplus y = x + y - xy, \tag{4.7}$$

$$x \bullet y = xy. \tag{4.8}$$

This particular algebraic object was selected to expand a Boolean model for computations using not only the Boolean numbers 0 and 1, but also probabilities of the occurrence of the Boolean 1. It is easy to see that function (4.5) corresponds to the conditional probability that the Boolean function $\overline{a}$ will take on the value of 1, if a takes on the value of 1 with probability x. Equation (4.6) corresponds to the conditional probability that the Boolean function (4.1) will take on the value of 1, if a takes on this value with the probability x and b with the probability y. Functions (4.7) and (4.8) can be interpreted similarly as the probabilities with which functions (4.3) and (4.4) will take on the value of 1.

Let us agree to write function (4.6) in exponential form:

$$F(x, y) = 1 - y + xy \stackrel{def}{=} x^y. \tag{4.9}$$

This is our special *convention* and it should not be confused with a regular operation of "raising to a power." Function $F(x, y)$, apparently, is connected with the fundamental basis of human cognition, because it appears not only in processes of choice, but also in the categorization of intensities of physical stimuli and even in the generation of musical intervals (see Lefebvre, 1992; 1997a).

5 The Subject, Free Will, and Freedom of Choice

response vs. conscious human act; intention; free will; freedom of choice; positive pole vs. negative pole; variables; axioms; deduction of the basic function; the Realist

Within the framework of our consideration, the agent can be represented by a hierarchy of images and relations. The process of computation, as was shown above, corresponds to cognitive process generating the agent's action. This action, however, resembles more closely a response than a conscious human act.

Real human actions are *intentional* and in some degree related to what we customarily call *free will* and *freedom of choice*. We can call an agent the subject only if he has these traits. Therefore our next task is to introduce the concepts of *free will* and *freedom of choice* into the model of the agent. To do this, we begin by constructing a model of the subject based on a few simple axioms that are not connected with the assumption that the subject has a reflexive structure. Then we demonstrate that it automatically follows from these axioms that the subject possesses the reflexive structure, and this allows us to find the element of the reflexive structure functioning as a "generator" of free will.

In beginning to construct our model of the subject we assume that he faces a choice between two alternatives. One plays the role of the positive pole, the other that of the negative. We presume that the subject corresponds to a function of three arguments:

$$X_1 = f(x_1, x_2, x_3), \tag{5.1}$$

where X_1, x_1, x_2, and x_3 are defined on the interval [0,1].

Variable x_1 shows the actual pressure of the world toward the positive pole at the *moment of choice.* Variable $\overline{x}_1 = 1 - x_1$ is the pressure toward the negative pole at the moment of choice. Variable x_2 represents the subject's sphere of knowledge. Its value is the world's pressure toward the positive pole as *expected* by the subject. This expectation can be based both on his own experience and on the available information. The value of $\overline{x}_2 = 1 - x_2$ is the pressure expected toward the negative pole. The term "expected" is used here in its direct meaning, not as a "mathematical expectation." Variable X_1 represents the subject's *executive sphere.* Its value is the probability with which the subject is *ready* to choose the positive pole. Variable x_3 represents the subject's *intentional sphere.* Its value is the probability with which the subject *wants* or *plans* to choose the positive pole. In this way we distinguish between the subject's real readiness to make a choice and his subjective desire or plan to do so.

To find function $X_1 = f(x_1, x_2, x_3)$, let us turn to our intuitive ideas about human beings, received from our culture and compounded by the work of many generations of philosophers and religious thinkers. The general concept of man in the context of choice can be formulated as four theses:

I. The subject has free will.

II. Under some circumstances, the subject has freedom of choice.

III. The external world can dictate that the subject choose "good."

IV. The external world can dictate that the subject choose "bad" only if the subject believes that the world is perfect.

Theses I and II formulate a basic distinction between the subject and the object. The subject can be a source and a reason for his own activity.

Theses III and IV show restrictions imposed on the subject by the external world. It can create the norms of behavior and make the subject choose what is considered "correct" (III). It also can incline the subject to choose what is considered "wrong" under the condition that the subject believes that the world is a source of absolute truth (IV).

The subject's free will is modeled by the assumption that variable x_3 takes on any value from the interval [0,1]. This assumption is not trivial. Let us imagine that for the function $X_1 = f(x_1, x_2, x_3)$ which we chose, there exists a pair x_1, x_2 for which exists a value $x_3 = c$ such that $X_1 < 0$. Since the probability cannot be negative, we have to conclude that the subject cannot have the intention $x_3 = c$, that is, he does not have free will. Therefore we have to reject this particular function, and the assumption about free will serves as a criterion for choosing function $X_1 = f(x_1, x_2, x_3)$ (Lefebvre, 1994; 1995; 1997b). The statement that, under certain circumstances, the subject has freedom of choice is modeled by the assumption that there exists at least one pair of values $x_1 = x_1^0$, $x_2 = x_2^0$ such that $X_1 \equiv x_3$; that is, any subject's intention (x_3) turns into readiness (X_1). In other words, with these values of pressure the subject can do what he wants.

In this way we distinguish between *free will* and *freedom of choice.* Free will is the freedom of desire; freedom of choice is the ability to turn any plan into reality. The subject always has free will, but the freedom of choice appears only under special circumstances.

The world's pressure toward "good" is modeled by assuming $x_1 = 1$, and toward "bad" by $x_1 = 0$. The words "the subject believes that the world is perfect" mean $x_2 = 1$; that is, the subject expects that the world will push him only toward "good." Thus, thesis III corresponds to the assumption that if x_1 is equal to 1, X_1 is also equal to 1, and thesis IV means that for $x_1 = 0$ and $x_2 = 1$ the value of X_1 is 0.

Let us introduce now one technical restriction on function $X_1 = f(x_1, x_2, x_3)$: we will presume that it is linear in respect to each variable. This restriction reflects our assumption that the function is rather simple.

The results of these considerations are summarized in the set of axioms: axioms 1-4 correspond to thesis I-IV, and axiom 5 to the technical restriction.

Axiom 1. Function $X_1 = f(x_1, x_2, x_3)$ is such that for any triple $x_1, x_2, x_3 \in [0,1]$, $X_1 \in [0,1]$.

Axiom 2. There exists a pair $x_1^0, x_2^0 \in [0,1]$ such that $f(x_1^0, x_2^0, x_3) \equiv x_3$, where x_3 is any number from [0,1].

Axiom 3. $f(1, x_2, x_3) \equiv 1$, where x_2 and x_3 are any numbers from [0,1].

Axiom 4. $f(0, 1, x_3) \equiv 0$, where x_3 is any number from [0,1].

Axiom 5. If the values of any two variables out of the three, x_1, x_2, x_3, are fixed, function $X_1 = f(x_1, x_2, x_3)$ is linear on the third.

With the help of these Axioms we will prove that

$$X_1 = x_1 + (1 - x_1)(1 - x_2)x_3. \tag{5.2}$$

It follows from Axiom 5 that function X_1 can be presented as a tri-linear form:

$$\begin{gathered} X_1 = a_0 + a_1 x_1 \\ +a_2 x_2 + a_3 x_3 + a_4 x_1 x_2 + \\ a_5 x_1 x_3 + a_6 x_2 x_3 + a_7 x_1 x_2 x_3, \end{gathered} \tag{5.3}$$

where $a_0, ..., a_7$ are unknown real numbers. Using Axioms 3 and 4 we transform (5.3) into

$$\begin{gathered} X_1 = x_1 + a_0(1 - x_1)(1 - x_2) + \\ a_3(1 - x_1)(1 - x_2)x_3. \end{gathered} \tag{5.4}$$

It follows from Axiom 2 that

$$\begin{gathered} x_1^0 + a_0(1 - x_1^0)(1 - x_2^0) + \\ a_3(1 - x_1^0)(1 - x_2^0)x_3 \equiv x_3, \end{gathered} \tag{5.5}$$

which is possible only if the following two equations hold:

$$x_1^0 + a_0(1 - x_1^0)(1 - x_2^0) = 0, \tag{5.6}$$

$$a_3(1 - x_1^0)(1 - x_2^0) = 1. \tag{5.7}$$

It follows from these equations that $a_0 = 0$, $x_1^0 = 0$, and $x_2^0 \neq 1$; thus equation (5.7) can be rewritten as

$$a_3(1 - x_2^0) = 1 \tag{5.8}$$

which leads to

$$a_3 \geq 1. \tag{5.9}$$

Since $a_0 = 0$, function (5.4) can be presented as follows:

$$X_1 = x_1 + a_3(1 - x_1)(1 - x_2)x_3. \tag{5.10}$$

Suppose $a_3 > 1$; if we choose $x_1 = 0$, $x_2 = 0$, and $x_3 = 1$, then we have

$$X_1 = a_3 > 1, \tag{5.11}$$

which contradicts Axiom 1. Thus, as follows from (5.9), a_3 can take on only one value, $a_3 = 1$, and equation (5.2) is deduced.

It follows from (5.8) that $x_2^0 = 0$; this means that the subject has freedom of choice only if $x_1 = 0$ and $x_2 = 0$; that is, when the world inclines him to choose the negative pole and he knows this.

Other systems of axioms leading to function (5.2) (Lefebvre, 1992; 1997a; Schreider, 1994; Miller, 1997c) or to alternative functions also exist (Krylov, 1994; Miller, 1997a,b,c).

It is clear from (5.2) that if the values of x_1 and x_2 are fixed (where $x_1 + x_2 > 0$), not every value of x_3 leads to $X_1 = x_3$; that is, not every plan can be realized. Imagine now that, for any pair x_1, x_2, the subject is capable of generating a plan such that equation

$X_1 = x_3$ holds. We will call such a subject a *Realist*, because he generates only those plans that can be turned into reality. The Realist corresponds to the equation:

$$x = x_1 + (1 - x_1)(1 - x_2)x, \tag{5.12}$$

where values of x_1 and x_2 are fixed. By solving this equation in relation to x, for $x_1 + x_2 > 0$, we obtain:

$$x = \frac{x_1}{x_1 + x_2 - x_1 x_2}. \tag{5.13}$$

It follows from (5.12) that for $x_1 = 0$ and $x_2 = 0$, that is, for $x_1 + x_2 = 0$, variable x becomes independent. Therefore the function of readiness for Realist is:

$$x = \begin{cases} \frac{x_1}{x_1+x_2-x_1x_2}, & \text{if } x_1 + x_2 > 0, \\ \text{any} \in [0,1], & \text{if } x_1 + x_2 = 0. \end{cases} \tag{5.14}$$

The Realist's free choice is related to chaotic behavior. It follows from (5.14) that for $x_1+x_2 > 0$ the value of x is unambiguously determined. For $x_1 + x_2 = 0$ that determination disappears. Consider the set of all the points (x_1, x_2) belonging to a unit square. In every point of this square, except (0,0), function (5.14) is continuous. The point (0,0) is the point of chaotic behavior, because in its any vicinity we can find a point (x_1, x_2), where $x_1 + x_2 > 0$, such that function (5.13) takes on any value of x given in advance. Thus, in the vicinity of point (0,0), tiny fluctuations may lead to significant changes of x. Therefore, the Realist's ability to make free choices is formally connected with chaotic behavior and instability.

Logical and epistomological aspects of this model were discussed by Jack Adams-Webber (1987; 1995; 1997), William Batchelder (1987), Louis Kauffman (1990), Vladimir Krylov (1994), Lev Levitin (1987), Ernest McClain (1990), Dudley Miller (1997c), Anatol Rapoport (1982; 1990; 1995; 1996), Julius Schreider (1994; 1998b), James Townsend (1983; 1990), Harvey Wheeler (1987; 1990a,b), Robert Zajonc (1987) and Karl Popper (1992). A special issue of the *Journal of Social and Biological Structures* (1987, Vol.10) was devoted to discussion of the model, and the enlarged version of this issue was published as a book (Ed. Wheeler, 1990). A discussion on this version is now in progress in the electronic journal of American Psychological Association, PSYCOLOQUY.

6 Relation between the Realist's Equation and the Logistic Differential Equation

formal neural net; deduction of the Realist's equation

We will demonstrate now that equation (5.13) can be deduced from very simple formal assumptions concerning the functioning of neuron nets. Let the subject's decision be generated by large bloc of neurons which can be arbitrarily connected among themselves. Each neuron can be in one of two states, which will be called positive and negative; a neuron in the positive state will be called positive neuron, and one in the negative will be called negative neuron.

We assume that the subject's readiness to choose the positive pole is equal to the ratio of the number of positive neurons to the total number of neurons. Thus, if the total number of neurons is equal to 1, and the fraction of positive ones is x, then the subject is ready to choose the positive pole with the probability x, and the negative pole with the probability $1 - x$.

We presume that a neuron in a positive state cannot transfer into a negative state, but that one in a negative state may transfer into a positive state. The system's work consists of consecutive transitional phases, each of which lasts a unit of time. The next phase begins immediately after the previous transitional phase ends. During one phase short-term connections are established between every negative neuron and some one other neuron from the net. If the second neuron is positive, the negative neuron transfers into the positive state; otherwise it remains negative. Let the probability that a negative neuron will connect with any other neuron during the transitional phase be identical, so that the probability that a negative neuron will connect with a positive neuron is x; since the fraction of negative neurons is $1-x$, then the fraction of positive neurons increases by $x(1-x)$ during one transitional period. We will consider that the appearance of new connections is equally distributed in time. Therefore, for the period Δt, the fraction of positive neurons grows by

$$\Delta x = x(1-x)\Delta t. \tag{6.1}$$

Expression (6.1) can be rewritten as

$$\frac{\Delta x}{\Delta t} = x(1-x). \tag{6.2}$$

For $\Delta t \to 0$, we obtain the differential equation

$$\frac{dx}{dt} = x(1-x). \tag{6.3}$$

This equation is called *logistic* (see, for example, Mickens, 1990).

If the initial value is given as $x = x_1$, where $x_1 \in [0,1]$, then the exact solution for the equation (6.3) is

$$x = \frac{x_1}{x_1 + x_2 - x_1 x_2}, \tag{6.4}$$

where

$$x_2 = e^{-t}. \tag{6.5}$$

The initial value of x_1 can be considered to be the proportion of positive neurons at the initial moment. Within this framework, the neurodynamic process of forming the subject's uncertain state is predetermined by the values of parameters x_1 and t. A direct "impression" of the situation at the moment of choice predetermines the proportions of positive and negative neurons, and the duration of the system's work from the moment 0 to the moment $t*$ predetermines the value of the subject's expectation x_2.

7 The Subject's Reflexive Structure

main theorem; image of the self; Realist's image of the self

Our next step consists of "extracting" a reflexive structure from function (5.2).

Statement. Function

$$F(x,y) = 1 - y + xy \tag{7.1}$$

is the only solution of the functional equation

$$\begin{aligned}\Phi(x_1, \Phi(x_2, x_3)) &= x_1 \\ &+(1-x_1)(1-x_2)x_3,\end{aligned} \tag{7.2}$$

where x_1, x_2, x_3 takes on any number from $[0,1]$ and all values of $\Phi(x_2, x_3)$ belong to $[0,1]$. The proof is given in Lefebvre, 1992a.

It follows from this statement that function (5.2) can be represented as

$$X_1 = F(x_1, F(x_2, x_3)), \tag{7.3}$$

or, in exponential form (see 4.9), as

$$X_1 = x_1^{x_2^{x_3}}. \tag{7.4}$$

We interpret the possibility of representing function (5.2) as (7.4) as formal "evidence" that the subject has a reflexive structure in which expression

$$X_2 = x_2^{x_3} \tag{7.5}$$

corresponds to his image of the self, and

$$X_3 = x_3 \tag{7.6}$$

to the image of the self held by the image of the self (Lefebvre, 1996; 1997a).

In the framework of the model introduced in section 5, however, variable x_3 represents the subject's intention. Thus we have established a connection between intention and reflexive structure: *the image of the self held by the image of the self fulfills the function of a generator of intentions.*

Expression (7.4) for the Realist can be written as

$$x = x_1^{x_2^{x}} \tag{7.7}$$

which is equal to equation (5.12). The Realist's image of the self corresponds to the function

$$X_2 = x_2^{x} \tag{7.8}$$

where x is given by equation (5.14). After transformation we obtain the following function:

$$X_2 = \begin{cases} \frac{x_2}{x_1+x_2-x_1x_2}, & \text{if } x_1 + x_2 > 0, \\ 1 - x, & \text{if } x_1 + x_2 = 0. \end{cases} \tag{7.9}$$

It follows from (5.14) and (7.9) that, for $x_1 = x_2 \neq 0$, the subject has an adequate image of the self, $(X_2 = x)$.

8 Two Aspects of Choice

an example of choice; pressure toward the poles; deontological aspect; utilitarian aspect

Let us begin by analyzing an example. Someone wants to sell a fake painting. If he tells the truth, he will receive \$200; if he conceals the truth, he will receive \$20,000. We will try to model the seller's choice by assuming that he is a Realist. First we have to single out the alternatives and relate them to the positive (+) and negative (-) poles. We assume that the seller shares traditional values:

to tell the truth (+) - to hide the truth (-)

Second, we have to correlate the sum of money that the seller can receive with the values of the variables. Suppose $x_1 = x_2$, that is, the pressure at the moment of choice is equal to the expected pressure; we assume that the value of the pressure toward each pole is proportional to the amount of money the seller would receive by choosing the corresponding alternative:

$$x_1 = x_2 = \frac{200}{200 + 20,000} = 0.0099...$$

By substituting this value into (5.14) we obtain

$$x = 0.502...$$

Therefore, given our assumptions, the seller will tell the truth with a probability slightly higher than 0.5.

This example illustrates how choice is schematized in reflexive models. We distinguish between two independent aspects of choice (Lefebvre, 1982; 1997a). The first one, which we call *deontological*, reflects the ability of the agent or the subject to accomplish a polarization of the alternatives. One of them plays the role of the positive pole (positive alternative), and the other that of the negative pole (negative alternative). The most noticeable example of polarization is the distinction between good and evil in moral choice.

The second aspect, *utilitarian*, reflects the difference between alternatives in their practical attractiveness for the subject. We assume that the choice is between the poles and that the alternative utility-measures predetermine the pressure toward poles.

Let us emphasize that, from the point of view of the deontological aspect, the alternatives are strictly polar, that is, they cannot have degrees of positivity or negativity. From the point of view of the utilitarian aspect, however, the alternatives can have various degrees of attractiveness to the subject, which can be expressed as a non-negative number. We connect the values of pressure, $x_1, \overline{x}_1$ and $x_2, \overline{x}_2$ with utilities of the alternatives by assuming that the pressure is proportional to utility-measures:

$$x_1 = \frac{v_1}{v_1 + v_2}, \qquad \overline{x}_1 = \frac{v_2}{v_1 + v_2}, \tag{8.1}$$

$$x_2 = \frac{u_1}{u_1 + u_2}, \qquad \overline{x}_2 = \frac{u_2}{u_1 + u_2}, \tag{8.2}$$

where $v_1, v_2 \geq 0$, $v_1 + v_2 > 0$ and $u_1, u_2 \geq 0$, $u_1 + u_2 > 0$. The numbers v_1 and v_2 are real utilities of the alternatives at the moment of choice; numbers u_1 and u_2 are the utility-measures formed before, beyond the specific conditions of a given particular problem.

By substituting x_1 and x_2 from (8.1) and (8.2) into equation (5.14) we obtain the following:

$$x = \begin{cases} \frac{v_1}{v_1 + \frac{u_1}{u_1+u_2} v_2}, \text{ if } \ v_1 + u_1 > 0, \\ \\ \text{any} \in [0,1], \text{ if } \ v_1 + u_1 = 0. \end{cases} \tag{8.3}$$

Therefore, the subject has freedom of choice only if $v_1 = 0$ and $u_1 = 0$, that is, only if he does not link the positive pole with any positive utility. This conclusion corresponds to a requirement present in all world religions: *good* must be separated from *utility* (see Lefebvre, 1997a).

9 Subject with a Model of the Situation

representation of situations by means of functions; bi-linear functions; list of sixteen equations

Let a subject be involved in a situation whose outcome depends both on his and on another subject's actions. He is able to play out in his mind various outcomes depending on his and the other's choice. As a result, the subject gains an idea of the probabilities that the situation will turn out to be positive or negative as a result of certain choices. Unlike the evaluation of alternatives, the evaluation of situations, from the point of view of the deontological aspect, is not strictly binary and can be expressed by a number from the interval [0,1].

The following reflexive structure corresponds to the subject who is playing out various situations in his mind:

$$\begin{array}{l} \qquad\qquad x_3 * y \\ \qquad x_2 \\ X_1 = x_1. \end{array} \tag{9.1}$$

Expression (9.1) is a generalization of expression (7.4). An image of the self at the subject described by equation (9.1) contains not only an image of the self (x_3), but also an image of the other (y) and of their relationship (*). Variable x_3, in addition, represents the subject's intention to choose the positive pole. Expression

$$x_3 * y = M(x_3, y) \tag{9.2}$$

will be called the subject's *model of the situation.* In general, any function whose arguments and values belong to the interval [0,1] can play the role of $M(x_3, y)$. We call the subject Realist, if for any $x_1, x_2, y \in [0,1]$ the subject's intention x_3 is equal to his readiness, that is, the following equation holds:

$$\begin{array}{l} \qquad M(x,y) \\ \quad x_2 \\ x_1 \qquad = x, \end{array} \tag{9.3}$$

or, in standard form:

$$x_1 + (1 - x_1)(1 - x_2)M(x, y) = x. \tag{9.4}$$

It is important to emphasize that equation (9.4) does not hold for every function $M = M(x, y)$, because only for some of them the equation has roots belonging to the interval [0,1].

Let $M(x_3, y)$ be a linear function in relation to each of the two variables, when the other variable's value is fixed, and let the value of M be integer when the values of arguments are integers (1 or 0) (Lefebvre, 1992; Miller, 1997a,b,c). It follows from the assumption of linearity that

$$M(x_3, y) = a_0 + a_1 x_3 + a_2 y + a_3 x_3 y. \tag{9.5}$$

The value $x_3 = 1$ means that the subject imagines himself having chosen the positive pole, and $x_3 = 0$ means that he imagines choosing the negative pole. When $y = 1$, he imagines the other subject choosing the positive pole, and when $y = 0$, the negative pole. We assume that the subject is capable of evaluating the situation as good or bad for any combination of integer variables. Thus we obtain a system of linear equations:

$$\begin{array}{l} M(1,1) = \alpha_1, \\ M(1,0) = \alpha_2, \\ M(0,1) = \alpha_3, \\ M(0,0) = \alpha_4, \end{array} \tag{9.6}$$

where α_i is either 1 or 0. By solving this system we find a function M, and if $X_1 = x_3$ we obtain an equation of the type (9.4). With the integer values of x_3 and y, function $M(x_3, y)$ can be considered a Boolean function. It is easy to see that there are only 16 different functions M. Below is a list of these functions, their limit values in a tabular and logical forms, canonical forms of equation (9.4) for each function, and solutions for equation (9.4).

Table 1: First M-Function

x_3	y	M_1
1	1	0
1	0	0
0	1	0
0	0	0

$M_1 \equiv 0$; Logical form: $M_1 \equiv 0$

Canonical form:

$$x - x_1 = 0$$

Solution:

$$x = x_1$$

Table 2: Second M-Function

x_3	y	M_2
1	1	1
1	0	1
0	1	0
0	0	0

$M_2 = x$; Logical form: $M_2 = x_3$

Canonical form:

$$(x_1 + x_2 - x_1 x_2)x - x_1 = 0$$

Solution:

$$x = \frac{x_1}{x_1 + x_2 - x_1x_2}, \text{if } x_1 + x_2 > 0,$$

if $x_1 + x_2 = 0$, any number from [0,1]

Table 3: Third M-Function

x_3	y	M_3
1	1	0
1	0	0
0	1	1
0	0	1

$$M_3 = 1 - x; \text{ Logical form: } M_3 = \overline{x}_3$$

Canonical form:

$$(1 + (1 - x_1)(1 - x_2))x$$
$$-(1 - x_2 + x_1x_2) = 0$$

Solution:

$$x = \frac{1 - x_2 + x_1x_2}{1 + (1 - x_1)(1 - x_2)}$$

Table 4: Fourth M-Function

x_3	y	M_4
1	1	1
1	0	0
0	1	1
0	0	0

$$M_4 = y; \text{ Logical form: } M_4 = y$$

Canonical form:

$$x - (1 - x_1)(1 - x_2)y - x_1 = 0$$

Solution:

$$x = x_1 + (1 - x_1)(1 - x_2)y$$

Table 5: Fifth M-Function

x_3	y	M_5
1	1	0
1	0	1
0	1	0
0	0	1

$$M_5 = 1 - y; \text{ Logical form: } M_5 = \overline{y}$$

Canonical form:

$$x + (1 - x_1)(1 - x_2)y$$
$$-x_1 - (1 - x_1)(1 - x_2) = 0$$

Solution:

$$x = x_1 + (1 - x_1)(1 - x_2)(1 - y)$$

Table 6: Sixth M-Function

x_3	y	M_6
1	1	1
1	0	0
0	1	0
0	0	0

$$M_6 = xy; \text{ Logical form: } M_6 = x_3\&y$$

Canonical form:

$$(1 - x_1)(1 - x_2)xy - x + x_1 = 0$$

Solution:
if $y \neq 1$, and $x_1 + x_2 \neq 0$,

$$x = \frac{x_1}{1 - (1 - x_1)(1 - x_2)y};$$

if $y = 1, x_1 + x_2 = 0$, any number from [0,1]

Table 7: Seventh M-Function

x_3	y	M_7
1	1	0
1	0	0
0	1	1
0	0	0

$M_7 = (1-x)y$; Logical form: $M_7 = \overline{x}_3 \& y$

Canonical form:

$$\begin{gathered}(1-x_1)(1-x_2)xy + x \\ -(1-x_1)(1-x_2)y - x_1 = 0\end{gathered}$$

Solution:

$$x = \frac{x_1 + (1-x_1)(1-x_2)y}{1 + (1-x_1)(1-x_2)y}$$

Table 8: Eight M-Function

x_3	y	M_8
1	1	0
1	0	1
0	1	0
0	0	0

$M_8 = x(1-y)$; Logical form: $M_8 = x_3 \& \overline{y}$

Canonical form:

$$\begin{gathered}(1-x_1)(1-x_2)xy + \\ (x_1 + x_2 - x_1x_2)x - x_1 = 0\end{gathered}$$

Solution:
if $y \neq 0$, and $x_1 + x_2 > 0$,

$$x = \frac{x_1}{(x_1 + x_2 - x_1x_2) + (1-x_1)(1-x_2)y};$$

if $y = 0$ and $x_1 + x_2 = 0$, any number from [0,1].

Table 9: Ninth M-Function

x_3	y	M_9
1	1	0
1	0	0
0	1	0
0	0	1

$M_9 = (1-x)(1-y)$; Logical form: $M_9 = \overline{x}_3 \& \overline{y}$

Canonical form:

$$\begin{gathered}(1-x_1)(1-x_2)xy \\ -(1+(1-x_1)(1-x_2))x \\ -(1-x_1)(1-x_2)y \\ +x_1 + (1-x_1)(1-x_2) = 0\end{gathered}$$

Solution:

$$x = \frac{1 - x_2(1-x_1) - (1-x_1)(1-x_2)y}{1 + (1-x_1)(1-x_2)(1-y)}$$

Table 10: Tenth M-Function

x_3	y	M_10
1	1	1
1	0	1
0	1	1
0	0	0

$M_{10} = x + y - xy$; Logical form: $M_{10} = x_3 \vee y$

Canonical form:

$$\begin{gathered}(1-x_1)(1-x_2)xy \\ +(1-(1-x_1)(1-x_2))x \\ -(1-x_1)(1-x_2)y - x_1 = 0\end{gathered}$$

Solution:
if $y \neq 0$ and $x_1 + x_2 > 0$,

$$x = \frac{x_1 + (1-x_1)(1-x_2)y}{1 - (1-x_1)(1-x_2) + (1-x_1)(1-x_2)y};$$

if $y = 0$ and $x_1 + x_2 = 0$, any number from [0,1]

Table 11: Eleventh M-Function

x_3	y	M_{11}
1	1	1
1	0	1
0	1	0
0	0	1

$M_{11} = 1 - y + xy$; Logical form: $M_{11} = x_3 \vee \overline{y}$

Canonical form:

$$(1 - x_1)(1 - x_2)xy - (1 - x_1)(1 - x_2)y + \\ x_1 - x + (1 - x_1)(1 - x_2) = 0$$

Solution:
if $y \neq 1$ and $x_1 + x_2 > 0$,

$$x = \frac{x_1 + (1 - x_1)(1 - x_2)(1 - y)}{1 - (1 - x_1)(1 - x_2)y};$$

if $y = 1$ and $x_1 + x_2 = 0$, any number from [0,1]

Table 12: Twelfth M-Function

x_3	y	M_{12}
1	1	1
1	0	0
0	1	1
0	0	1

$M_{12} = 1 - x + xy$; Logical form: $M_{12} = \overline{x}_3 \vee y$

Canonical form:

$$(1 - x_1)(1 - x_2)xy + x_1 + (1 - x_1)(1 - x_2) \\ -(1 + (1 - x_1)(1 - x_2))x = 0$$

Solution:

$$x = \frac{1 - x_2 + x_1x_2}{1 + (1 - x_1)(1 - x_2)(1 - y)}$$

Table 13: Thirteenth M-Function

x_3	y	M_{13}
1	1	0
1	0	1
0	1	1
0	0	1

$M_{13} = 1 - xy$; Logical form: $M_{13} = \overline{x}_3 \vee \overline{y}$

Canonical form:

$$(1 - x_1)(1 - x_2)xy + x \\ -x_1 - (1 - x_1)(1 - x_2) = 0$$

Solution:

$$x = \frac{1 - x_2 + x_1x_2}{1 + (1 - x_1)(1 - x_2)y}$$

Table 14: Fourteenth M-Function

x_3	y	M_{14}
1	1	0
1	0	1
0	1	1
0	0	0

$M_{14} = x + y - 2xy$
Logical form: $M_{14} = (\overline{x}_3 \& y) \vee (x_3 \& \overline{y})$

Canonical form:

$$2(1 - x_1)(1 - x_2)xy + (x_1 + x_2 - x_1x_2)x - \\ (1 - x_1)(1 - x_2)y - x_1 = 0$$

Solution:
if $y \neq 0$ and $x_1 + x_2 > 0$,

$$x = \frac{x_1 + (1-x_1)(1-x_2)y}{(x_1 + x_2 - x_1x_2) + 2(1-x_1)(1-x_2)y};$$

if $y = 0, x_1 + x_2 = 0$, any number from $[0, 1]$

Table 15: Fifteenth M-Function

x_3	y	M_{15}
1	1	1
1	0	0
0	1	0
0	0	1

$$M_{15} = 1 - x - y + 2xy;$$

Logical form: $M_{15} = (\overline{x}_3 \vee y)\&(x_3 \vee \overline{y})$

Canonical form:

$$\begin{gathered} 2(1-x_1)(1-x_2)xy \\ -(1 + (1-x_1)(1-x_2))x \\ -(1-x_1)(1-x_2)y + 1 - x_2 + x_1x_2 = 0 \end{gathered}$$

Solution: if $y \neq 1, x_1 + x > 0$,

$$x = \frac{1 - x_2(1-x_1) - (1-x_1)(1-x_2)y}{1 + (1-x_1)(1-x_2)(1-2y)};$$

if $y = 1, x_1 + x_2 = 0$, any number from [0,1]

Table 16: Sixteenth M-Function

x_3	y	M_{16}
1	1	1
1	0	1
0	1	1
0	0	1

$M_{16} \equiv 1$; Logical form: $M_{16} \equiv 1$

Canonical form: $x + x_2 - x_1x_2 - 1 = 0$
Solution: $x = 1 - x_2 + x_1x_2$

Each of the solutions given above can be represented as a function of y depending on the parameters x_1 and x_2:

$$x = P^i_{x_1,x_2}(y), \tag{9.7}$$

where $i = 1, 2, ..., 16$ and y can take on any value from [0,1]. With any set of four i, x_1, x_2, y, there exists a value of x, so that the subject modeling the situation with the functions given above can be called Realist. Note that not all the functions of the set (9.7) are single-valued, but for any pair x_1 and x_2 each function contains a branch continuous on the interval [0,1]. Every case for which a function of the type (9.7) is not single-valued corresponds to the condition where the subject has freedom of choice.

10 Examples of Modeling the Process of Choice

analysis of four situations

Imagine the following situation. Two soldiers are captured by the enemy, who knows that they possess important information, tries to force each soldier to collaborate. We assume that for both soldiers the alternative "not to collaborate" plays the role of the positive pole, and "to collaborate" that of the negative pole. We also assume that the pressure at the moment of interrogation is so high that $x_1 = 0$, and since the soldiers knew in advance what to expect in captivity, $x_2 = 0$ as well. In the framework of this scenario, we will analyze four cases.

Case 1. The interrogators inform the soldiers that they will survive if at least one of them collaborates.

Let us designate one soldier **X** and the other **Y** and model the behavior of soldier

X. First, let us find function $M(x,y)$ corresponding to the soldier **X**. We assume that the outcome of the situation can be considered positive for him only if he survives, and he can survive only if at least one of the soldiers collaborates. The following table corresponds to the model of the situation:

Table 17: Captured Soldiers Survive Only When At Least One Collaborates

x_3	y	M
1	1	0
1	0	1
0	1	1
0	0	1

The value of y in this table is **Y**'s choice as it is imagined by **X**. By referring to the list of functions in section 9 we find that:

$$M = M_{13} = 1 - xy. \qquad (10.1)$$

Since we assumed that $x_1 = 0$ and $x_2 = 0$, soldier **X**, if he is a Realist, corresponds to equation

$$\begin{matrix} & & 1 - xy & \\ & 0 & & \\ 0 & & & = x, \end{matrix} \qquad (10.2)$$

which in the standard form looks as follows:

$$1 - xy = x. \qquad (10.3)$$

We find from (10.3) that

$$x = \frac{1}{1+y}. \qquad (10.4)$$

A graph of function (10.4) is represented in Fig.4. Let $y = 0$. This means that **X** has learned that **Y** agreed to collaborate. At this moment **X**'s model of the situation changes to the state $M_{13}(x_3, 0) \equiv 1$. We interpret this as **X**'s awareness that his life is saved; then $x = 1$, that is, **X** makes the decision not to collaborate. The choice deduced from the model corresponds to our intuition: having learned that his life is saved, soldier **X** easily rejects collaboration with the enemy.

Let $y = 1$. This means that **X** learned that **Y** refused to collaborate. **X**'s model of the situation changes to the state $M_{13}(x_3, 1) = 1 - x_3$. We interpret this as **X**'s feeling that his life is in his own hands. If he decides not to collaborate, then $x_3 = 1$, $M_{13}(1,1) = 0$ and **X** understands that he will be killed. If he decides to collaborate, then $x_3 = 0$, $M_{13}(0,1) = 1$ and **X** understands that his life is saved.

Looking at the graph on Fig.4, we see that for $y = 1$, $x = (1/2)$, which corresponds to the soldier's transfer into an uncertain state; being there he chooses between the alternatives with an equal probability. In this case the soldier is not as firm as in the first case, where he learns that the other soldier has collaborated; this also confirms our intuition.

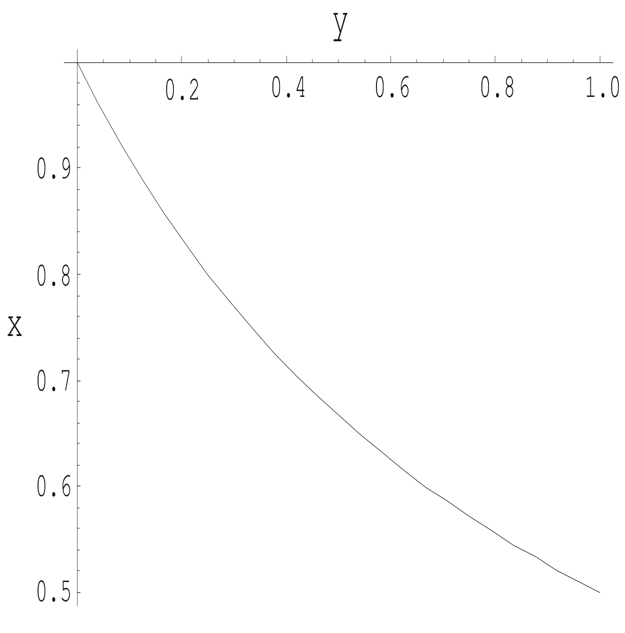

Figure 4: Graph of the soldier's readiness not to collaborate with the enemy under the condition that both he and his friend will survive if at least one of them collaborates.

The graph shows that as **Y**'s implacability increases (y grows), **X**'s implacability decreases (x diminishes), but his readiness to collaborate with the enemy never exceeds $(1/2)$ because $x \geq (1/2)$.

Case 2. The interrogators say that the soldiers can survive only if both collaborate. Then **X**'s model of the situation corresponds to the following table:

Table 18: Captured Soldiers Survive Only When Both Collaborates

x_3	y	M
1	1	0
1	0	0
0	1	0
0	0	1

From this table we find that

$$\begin{aligned} M &= M_9 = (1-x)(1-y) \\ &= 1 - x - y + xy. \end{aligned} \tag{10.5}$$

Subject **X** is described by equation

$$\begin{array}{l} 1 - x - y + xy \\ 0 \\ 0 \qquad\qquad\qquad = x, \end{array} \tag{10.6}$$

or

$$1 - x - y + xy = x, \tag{10.7}$$

and

$$x = \frac{1-y}{2-y}. \tag{10.8}$$

Let $y = 0$. This means that **X** learns that the other soldier has agreed to collaborate. At this moment **X**'s readiness to collaborate becomes equal to $(1/2)$ as shown in Fig.5.

Let $y = 1$; **X** learns that **Y** has rejected collaboration. Then **X**'s model of the situation changes to the state $M_9(x_3, 1) \equiv 0$, which

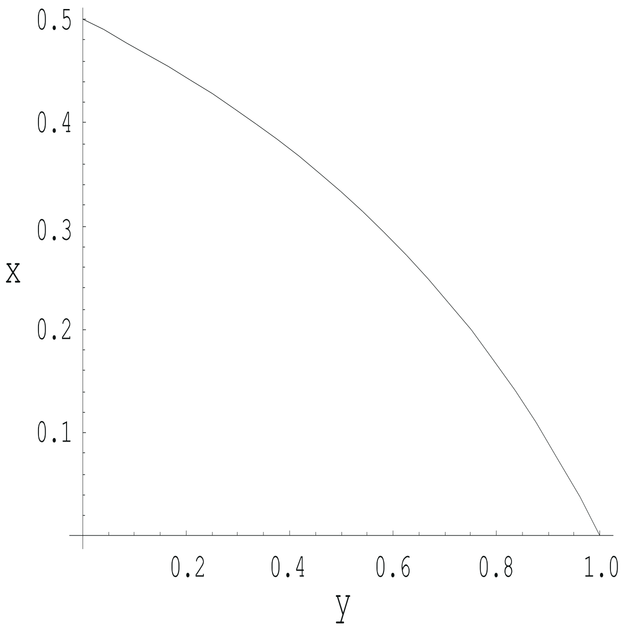

Figure 5: Graph of the soldier's readiness not to collaborate with the enemy under the condition that both he and his friend will survive only if both collaborate.

we interpret as **X**'s awareness of the fact that he will be killed. At this moment, x becomes equal to 0, that is, the soldier wants to collaborate with the enemy. When y grows, x decreases. Thus, the less **Y** is ready to collaborate, the more **X** agrees to do so.

Case 3. The soldiers come to the conclusion that they can survive if they both behave in the same way: if both are silent, the enemy will decide they do not have important information; if both collaborate, the enemy will reward them by letting them live. This case corresponds to Table 19.

Table 19: Captured Soldiers Assume Survival Only When Both Behave Equivalently

x_3	y	M
1	1	1
1	0	0
0	1	0
0	0	1

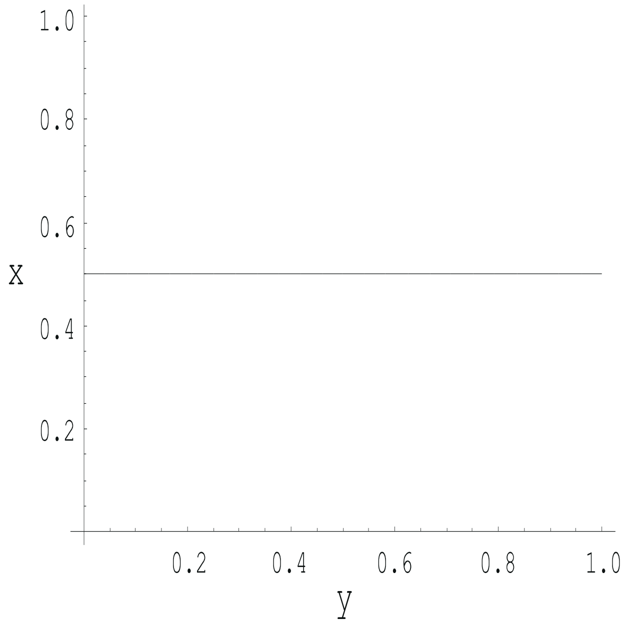

Figure 6: Graph of the soldier's readiness under the condition that they will both survive if they behave in the same way.

From this table we find that $M = M_{15}$ and that **X** is described by the equation

$$\begin{array}{l} 1-x-y+2xy \\ 0 \\ 0 \qquad\qquad\quad = x, \end{array} \tag{10.9}$$

or

$$1-x-y+2xy = x; \tag{10.10}$$

this implies:

$$x = \begin{cases} 1/2, \text{ if } \ y < 1, \\ \text{any} \in [0,1], \text{ if } \ y = 1. \end{cases} \tag{10.11}$$

Let $y < 1$; this means that **X** learns that **Y**, in principle, may compromise with the enemy. In this case, independently of the value of y, **X** will choose between the alternatives with an equal probability (see Fig.6).

Let $y = 1$; **X** learns that **Y** has rejected compromise. In this case **X** obtains the ability to make a free choice. The probability of his choosing any alternative is determined by him, and therefore cannot be predicted.

Case 4. Each soldier understands that he can survive only if he keeps silence, because only then will he continue to be valuable for the enemy as a potential informer. This case corresponds to Table 20.

Table 20: Captured Soldier Survives Only When Silence Maintained

x_3	y	M
1	1	1
1	0	1
0	1	0
0	0	0

From this table we find that $M = M_2$ and **X** is described by equation

$$\begin{array}{l} x \\ 0 \\ 0 \quad = x, \end{array} \tag{10.12}$$

which is equal to the identity

$$x \equiv x, \tag{10.13}$$

where x is any number from [0,1]. We see that in this case **X**'s readiness to reject collaboration does not depend on his friend's state. The fact that x can take on any value from [0,1] means that **X** has freedom of choice.

11 Surprise Actions

unlikely events; decisions in surprise situations

Surprise results from the occurrence of an event for which our expectation is quite low. In the reflexive model, expectation is presented by the variable x_2, and reality at the moment of choice by the variable x_1. We consider that the more x_1 deviates from x_2, the more unexpected for the subject is the world's influence x_1.

Consider the following example. Soldier **X** has been called in for multiple interrogations, but was treated gently and even told that he would not be shot under any circumstances. As a result, **X** develops a model of the situation $M_{16} \equiv 1$. Suppose that after a series of interrogations, the values of x_1 and x_2 became equal to $(4/5)$. Now the soldier corresponds to equation

$$(4/5)^{(4/5)^{M_{16}(x,y)}} = x, \qquad (11.1)$$

or

$$(4/5)^{(4/5)^{1}} = x, \qquad (11.2)$$

whence

$$x = (21/25) = 0.84. \qquad (11.3)$$

In this state, the soldier is ready to reveal secrets to the enemy with a probability of 0.16.

Suppose that now the soldier is told that he will be shot immediately if he does not reveal all the information he possesses. At that moment the soldier's expectation of the positive pressure had been $x_2 = (4/5)$, but in reality it turns out to be $x_1 = 0$. The threat changes his model of the situation: function $M_{16} \equiv 1$ changes to $M_3 = 1 - x$ and the equation corresponding to the soldier is now

$$0^{(4/5)^{1-x}} = x. \qquad (11.4)$$

The root of this equation is

$$x = (1/6) = 0.166... \qquad (11.5)$$

Therefore in this situation the soldier is ready for collaboration with the enemy with a probability of 0.834...

We see that the high expectation of a positive outcome weakens the soldier's ability to withstand strong negative influence. Note that the following identity holds:

$$x_1^{1^{M(x_3,y)}} \equiv x_1, \qquad (11.6)$$

where M is an arbitrary function taking on values from [0,1]. The identity (11.6) means that when the subject does not expect pressure toward the negative pole ($x_2 = 1$) he is completely vulnerable to pressure applied to him at the moment of choice ($x = x_1$), so that the surprising appearance of the value $x_1 = 0$ determines the subject's choice of the negative pole.

12 Ideal Behavior

heroic behavior; equation related to heroic behavior

We see that soldier **X** in captivity does not demonstrate high spirit and, in one way or another, makes his moral choice depending on circumstances. We all know, however, examples of the truly heroic behavior of people who remain loyal to their moral principles even when facing a choice between life and death. We will analyze these cases as well.

Let us pose the question: what factors influence the Realist's behavior? There are two groups of factors. The first group of factors determine the values of x_1 and x_2, and the second ones generate the function $M(x, y)$. The ideal subject is capable of resisting the influence of these factors. At the end of section 8 we refer to a subject who does not link the positive pole with positive utility. For this subject, $x_1 \equiv 0$ and $x_2 \equiv 0$, independently of the external world's (or enemy's) influence. But this is not enough for ideal behavior. In order to always choose the positive pole, the subject must always schematize the situation with the same function $M_{16} \equiv 1$, corresponding to the following table:

Table 21: Ideal Behavior

x_3	y	M_{16}
1	1	1
1	0	1
0	1	1
0	0	1

There are only "ones" in the right hand column of this table; that is, the subject considers any outcome to be positive, including those leading to his death; this means that, for him, death is an *acceptable* outcome. The following equation always holds for this subject:

$$\begin{array}{l} \quad 1 \\ \;0 \\ 0 \quad = 1. \end{array} \tag{12.1}$$

Independently of the outer world's influence, this subject chooses the positive pole.

13 Imitation of the Other

imitation vs. reflexion; systems of equations; examples

Let us return to the examples analyzed in section 10. We assumed there that prior to making any decision the subject receives information about the other's readiness to choose the positive pole. Imagine now that **X** is isolated from **Y** and cannot receive information about the value of y, so that his cognitive system cannot compute the value of x. What do real people do in such a situation? They put themselves into the other's shoes and try to imitate the other's behavior. Such an imitation differs from reflexion. Reflexion is mental representation of the self and of the other with multiple mutual representations. Imitation is not a mental representation; the subject employs the entire self for modeling the other; he activates not only his cognitive sphere, but his motor sphere as well. He resembles an actor, playing the role of another person (see Fink, 1980). Imitation is an act of metamorphosis. This "double existence" is well-known to people in show business and to psychologists who study military thinking and styles of chess play.

From the viewpoint of our model, the process of imitating **Y** by **X** will be represented by a system of two equations:

$$\begin{array}{l} \quad M_X(x,y) \\ \;x_2 \\ x_1 \qquad = x \\ \\ \quad M_Y(y,x) \\ \;y_2 \\ y_1 \qquad = y. \end{array} \tag{13.1}$$

The first equation corresponds to subject **X** himself, the second to **X** playing the role of **Y**. We consider **X**'s decision-making process as a solution for the system (13.1). In standard form, it is a system of the type

$$\begin{array}{l} a_1xy + a_2x + a_3y + a_4 = 0 \\ \\ b_1yx + b_2y + b_3x + b_4 = 0, \end{array} \tag{13.2}$$

where the values of $a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4$ depend on x_1, x_2, y_1, and y_2. Each of these equations can be written in the form of (9.7) and system (13.2) presented as follows:

$$\begin{array}{l} x = P_1(y) \\ \\ y = P_2(x), \end{array} \tag{13.3}$$

where each function has a continuous branch on the interval [0,1]. By using the theorem of the fixed point, we find that system (13.3) has at least one solution in the area $0 \leq x,\ y \leq 1$. We will now analyze situations 1-4 as described in section 10, with the added condition that there is no contact between the captured soldiers, and therefore that soldier **X**'s imitational-cognitive system generates his readiness by solving system of equations of the type (13.1).

Case 1. Soldier **X** imitates soldier **Y** with the same function $M_{13} = 1-xy$ which he uses for himself. Therefore, he corresponds to the system of equations:

$$\begin{array}{l} 1-xy \\ 0 \\ 0 \quad = x \\ \\ 1-yx \\ 0 \\ 0 \quad = y, \end{array} \tag{13.4}$$

or, in standard form:

$$\begin{aligned} 1-xy &= x \\ 1-yx &= y. \end{aligned} \tag{13.5}$$

It follows from (13.5) that $x = y$ and the value of x must satisfy equation

$$x^2 + x - 1 = 0, \tag{13.6}$$

whose solution is

$$x = \frac{\sqrt{5}-1}{2} = 0.618... \tag{13.7}$$

Thus this soldier will choose the positive pole with a probability equal to the value of the golden section.

Case 2. The function $M_9 = 1-x-y+xy$ is used:

$$\begin{array}{l} 1-x-y+xy \\ 0 \\ 0 \quad = x \\ \\ 1-x-y+xy \\ 0 \\ 0 \quad = y, \end{array} \tag{13.8}$$

or

$$\begin{aligned} 1-x-y+xy &= x \\ 1-x-y+yx &= y. \end{aligned} \tag{13.9}$$

It follows from (13.9) that $x = y$ and the value of x will satisfy the equation

$$x^2 - 3x + 1 = 0, \tag{13.10}$$

and

$$x = 1 - \frac{\sqrt{5}-1}{2} = 0.382...$$

This is the probability with which soldier **X** will reject collaboration with the enemy; therefore, he will agree to collaborate with a probability equal to the golden section value.

Case 3. Function $M_{15} = 1-x-y+2xy$ is used:

$$\begin{array}{l} 1-x-y+2xy \\ 0 \\ 0 \quad = x \\ \\ 1-x-y+2xy \\ 0 \\ 0 \quad = y, \end{array} \tag{13.11}$$

or

$$\begin{aligned} 1-x-y+2xy &= x \\ 1-x-y+2yx &= y. \end{aligned} \tag{13.12}$$

It follows from (13.12) that $x = y$ and

$$x^2 - \frac{3}{2}x + \frac{1}{2} = 0. \tag{13.13}$$

Both roots of this equation belong to the interval [0,1]: $x^{(1)} = 1$, $x^{(2)} = (1/2)$. Therefore the system (13.11) has two solutions:

$$x^{(1)} = 1, \qquad y^{(1)} = 1$$
$$x^{(2)} = (1/2), \qquad y^{(2)} = (1/2). \tag{13.14}$$

We consider the process of solving the systems as a theoretical analogue for the imitational-cognitive process which generates the values of readiness, x. It follows from equations (13.14) that, in this specific case, the process cannot be completed without additional information about the value of y. In analyzing system (13.11) the subject's cognitive system concludes that if $y = 1$, then the subject will be transferred into a state with readiness $x = 1$, and if $y = (1/2)$, then to the state with readiness $x = (1/2)$. The subject's cognitive system "does not know," however, what **Y**'s state of readiness is in reality, so it comes to a dead end, which is registered by the subject as a sharp desire to obtain new information.

Imagine that soldier **X** looks through a window and sees soldier **Y** leaving interrogation; it is clear from **Y**'s face expression that he did not speak. At this moment **X**'s cognitive system emerges from the dead end, variable y takes on the value of 1, and the subject transfers to the state of readiness $x = 1$, that is, he makes the decision to remain silent.

Case 4. The subject uses function $M_2 = x$, when he "plays the self" and equivalent function $M_2 = y$, when he "plays the other:"

$$\begin{array}{lll} & & x \\ & 0 & \\ 0 & & = x \\ \\ & & y \\ & 0 & \\ 0 & & = y. \end{array} \tag{13.15}$$

Any pair x, y in [0,1] is a solution of this system, that is, soldier **X** has freedom of choice independently of any decision he made while playing the role of the other.

14 Subject with a Quadratic Model of the Situation

dilemma; deduction of equation; solution; index of frustration

In the previous sections we supposed that the subject has a bi-linear model of the situation. Let us now consider an example of a subject with a quadratic model. We will analyze a simple situation in which the subject makes his decision independently of his partners and he still can conduct a deontological assessment of the consequences of his choice.

Imagine that three soldiers **X**, **Y**, and **Z** are captured. The enemy exerts strong pressure on **X** to provide information. Soldier **X** knows that any choice he makes will become a death sentence for one of his comrades: if he is silent, **Y** will die; if he speaks, **Z** will die. We put **X** into correspondence to the following equation:

$$\begin{array}{lll} & & m(x_3) \\ & x_2 & \\ x_1 & & = X_1, \end{array} \tag{14.1}$$

where $m(x_3)$ is a square form:

$$m(x_3) = bx_3^2 + ax_3 + c \tag{14.2}$$

and a, b, c are real numbers.

The conditions of this particular problem allow us to find the value of c and to establish a relation between a and b. The soldier knows that regardless of his choice the result will be negative. We can represent his knowledge by the following table:

Table 22: Fatalist Behavior

x_3	$m(x_3)$
0	0
1	0

By substituting the values from this table into (14.2) we obtain:

$$a + b + c = 0$$
$$c = 0. \tag{14.3}$$

Whence that $b = -a$ and

$$m(x_3) = ax_3(1 - x_3). \tag{14.4}$$

It follows from inequality

$$0 \leq m(x_3) \leq 1 \tag{14.5}$$

that

$$0 \leq a \leq 4. \tag{14.6}$$

For the Realist ($X_1 = x_3 = x$) equation (14.1) can be rewritten:

$$\begin{gathered} a(1 - x_1)(1 - x_2)x^2 + \\ (1 - a(1 - x_1)(1 - x_2))x - x_1 = 0. \end{gathered} \tag{14.7}$$

The roots of this equation which belong to the interval [0,1] are given by the following expression (wherein $B = (1 - x_1)(1 - x_2)$, $\bar{Ba} = 1 - Ba$, $\bar{x_2} = 1 - x_2$, and $\bar{x_1} = 1 - x_1$):

$$x = \begin{cases} \frac{\sqrt{\bar{Ba}^2 + 4Bax_1} - \bar{Ba}}{2Ba}, x_1 > 0, \\ \left. \begin{matrix} 0 \\ \frac{a\bar{x_2} - 1}{a\bar{x_2}}, \text{if } a\bar{x_2} > 1 \end{matrix} \right\}, x_1 = 0 \end{cases} \tag{14.8}$$

if $a > 0, x_1 \neq 1,$ and $x_2 \neq 1,$ else,

$$x = \begin{cases} 1, \text{ if } x_1 = 1 \\ x_1, \text{ if } x_2 = 1, \text{ or } a = 0. \end{cases}$$

Suppose that the pressure on soldier **X** is so great that we can consider $x_1 = 0$ and $x_2 = 0$. The following equation corresponds to **X**:

$$\begin{matrix} ax(1 - x) & \\ 0 & \\ 0 & = x, \end{matrix} \tag{14.9}$$

or, in standard form,

$$ax(1 - x) = x. \tag{14.10}$$

This equation has two roots:

$$x^{(1)} = 0, \tag{14.11}$$

$$x^{(2)} = \frac{a - 1}{a}, \tag{14.12}$$

where $0 < a \leq 4$. It is clear that $x^{(2)}$ belongs to the interval [0,1] only under condition that $a \geq 1$.

We begin consideration with the values $a < 1$. For these values, equation (14.9) has only one non-negative root, $x^{(1)} = 0$. This means that soldier **X** capitulates, and his friend **Y** dies. When $a = 1$, another non-negative root appears, $x^{(2)} = 0$. But this does not change **X**'s behavior, because both roots correspond to the same state, in which **X**'s readiness to choose the positive pole is equal to 0.

Consider now cases in which $a > 1$. In addition to $x^{(1)} = 0$, equation (14.9) has one more root not equal to zero; it is given by equation (14.12). Then the soldier's cognitive system cannot bring **X** to one particular fixed state of readiness by solving equation (14.9). The soldier's psychological state can be described as follows: "On the one hand, he is ready to answer the enemy's questions with a probability of 1; but on the other hand, he is ready to resist with a probability of $(a - 1)/a$." Therefore, the soldier has two contradicting plans of actions; the greater a, the stronger contradiction, which reaches its maximum at $a = 4$, when, on the one hand, the soldier is sure that he must capitulate, and on the other, he is ready to resist with the probability equal to 3/4. The soldier perceives such a situation as a harsh psychological drama, whose depth is reflected in the value of a. Let us call it the subject's *index of frustration.* How can we find its value? We must apply certain restrictions to function (14.4). The value $x_3 = (1/2)$

corresponds to that psychological state of the subject in which he consciously refuses to give preference to one alternative over the other one. In this case

$$m(1/2) = \frac{a}{4}. \tag{14.13}$$

The value of $(a/4)$ is the subject's estimation of the development of events such that the chances of survival for each of his friends are equal. If the subject estimates this situation positively, then $(a/4) = 1$ and $a = 4$; if he estimates it negatively, then $(a/4) = 0$ and $a = 0$. In general, if the subject estimates this situation with an arbitrary number α from the interval [0,1], then $(a/4) = \alpha$ and $a = 4\alpha$.

15 Streams of Consciousness and Difference Equations

iterations; convergence; oscillations; logistic difference equation; bifurcation; chaotic stream of consciousness

We have presumed that the subject's cognitive-imitational system is capable of solving equations and even systems of equations. How does it do this? What calculations does it carry out? In this section we will consider a hypothetical mechanism for such computations: we will assume that their essence consists of consequent iterations which can be represented with the help of difference equations and systems of difference equations (Lefebvre et al., 1969; Lefebvre, 1982).

We showed in section 7 that variable x_3 can play two roles: it can represent the subject's intention as well as his cognizant image of the self. Let the subject have the ability to modify this image. If at moment n the subject represented by equation (7.4) has the readiness $x_{(n)}$, then the following sequence will correspond to it:

$$\begin{array}{ll} \quad x_{(n)} & \\ \; x_2 & \\ x_1 & = x_{(n+1)}, \end{array} \tag{15.1}$$

where x_0 is an arbitrary number from [0,1]. It is easy to prove that this sequence converges to the value (5.14) at any values of $x_1, x_2, x_0 \in [0,1]$. Thus the psychological process corresponding to equation (15.1) goes as follows: the subject realizes his free will by choosing an initial value of x_0; after that begins the iterative process of modifying his cognizant image of the self. Using traditional psychological terminology we will call this process the *stream of consciousness* (see Pope & Singer, 1978). The result of this process is an iterative solution of the equation

$$\begin{array}{ll} \quad x & \\ \; x_2 & \\ x_1 & = x. \end{array} \tag{15.2}$$

The value of x is a fixed point of transformation (15.1).

Consider the system of equations (13.4). It corresponds to the following system of difference equations:

$$\begin{array}{ll} \quad 1 - x_{(n)}y_{(n)} & \\ \; 0 & \\ 0 & = x_{(n+1)} \\ & \\ \quad 1 - x_{(n)}y_{(n)} & \\ \; 0 & \\ 0 & = y_{(n+1)}, \end{array} \tag{15.3}$$

where $x_{(0)}$ and $y_{(0)}$ are numbers from [0,1].

In standard form system (15.3) looks as follows:

$$\begin{array}{l} 1 - x_{(n)}y_{(n)} = x_{(n+1)} \\ \\ 1 - x_{(n)}y_{(n)} = y_{(n+1)}. \end{array} \tag{15.4}$$

Since the left hand sides of the equations are the same, $x_{(n)} = y_{(n)}$ at $n \geq 1$, system (15.4) corresponds to the equation

$$1 - x_{(n)}^2 = x_{(n+1)}. \qquad (15.5)$$

Consider a case in which the values of x_0 and y_0 are such that their product, $x_0 y_0$, differs from 0 and 1; then

$$x_{(n)} \to \frac{\sqrt{5}-1}{2}, \qquad (15.6)$$

when $n \to \infty$. The iterative process (15.5) solves the system of equation (13.4).

Consider a case in which $x_0 y_0$ equals either 1 or 0. In this case equation (15.5) generates periodical oscillations of the stream of consciousness: the sequence of x_n consists of the alternating values 1 and 0. The subject experiences leaps of readiness. For example, being at step n, he is ready to choose the positive pole with probability 1 ($x_n = 1$), then at step $n+1$, he is ready to choose the negative pole with probability 1 ($x_{n+1} = 0$). Such streams of consciousness are not stable. If, for any small change of x_0 and y_0 their product takes on a value different from 1 or 0, the alternating appearance of 1 and 0 will stop, and the system will tend to a stable state corresponding to the fixed point $x = (\sqrt{5}-1)/2$. We will now return to the analysis of the subject's mental domain under conditions of frustration (see section 14). Equation (14.9) corresponds to the following difference equation:

$$\begin{array}{ll} \quad ax_{(n)}(1-x_{(n)}) & \\ \ 0 & \\ 0 & = x_{(n+1)}, \end{array} \qquad (15.7)$$

or, in standard form,

$$ax_{(n)}(1-x_{(n)}) = x_{(n+1)}, \qquad (15.8)$$

where $0 \leq a \leq 4$.

Equation (15.8) is well known in the theory of dynamic systems and is called *logistic* (May, 1976; see also Mickens, 1990). If $0 \leq a \leq 1$,

$$\lim_{n\to\infty} x_n = 0 \qquad (15.9)$$

for any $x_{(0)} \in [0,1]$. Thus, in this case, there is only one fixed point, $x = 0$, which is stable. If $1 < a < 3$, this fixed point is not stable any longer, but another fixed stable point appears:

$$x = \frac{a-1}{a}. \qquad (15.10)$$

When $a = 3$, this point looses its stability, and no other stable fixed points appear; instead, there appear stable cycles with the length of 2. With further increase of a, the avalanche process of bifurcation goes on, that is, the appearance of new stable cycles and disappearance of old ones (see Fig.7). The system's behavior depends significantly on the initial value of $x_{(0)}$.

Therefore, we see that with an increasing index of frustration, a, the subject's stream of consciousness becomes chaotic and unpredictable.

16 Streams of Consciousness and Acts of Awareness

mathematical representation of awareness; creating new fixed points; awareness as a means to eliminate chaos in the stream of consciousness

We will show in this section that the act of awareness can be understood as a special transformation of the computational process creating new fixed points in the streams of consciousness.

First we have to indicate a formal analogue for the act of awareness. From the intuitive point of view, the act of awareness

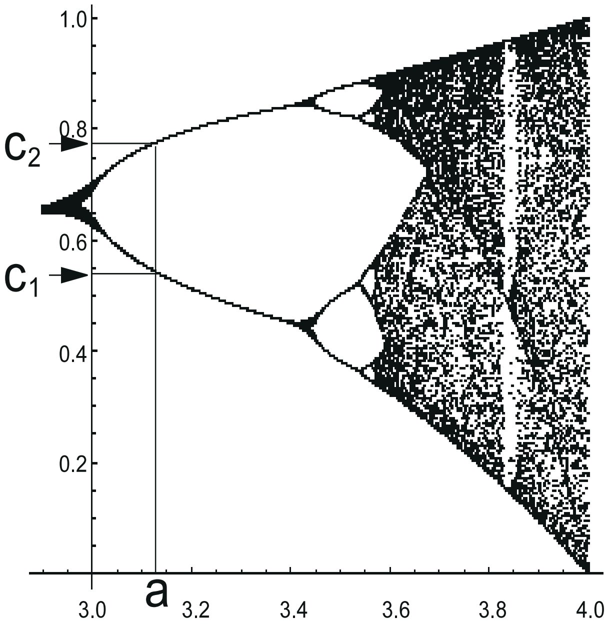

Figure 7: Bifurcation diagram for the logistic difference equation.

is a transformation of the subject in which the subject's reflexive structure is reflected into his "model of the self," and the subject obtains a new reflexive structure (Lefebvre, 1967; 1977a,b; 1985).

At a particular moment the subject has the following reflexive structure:

$$X_1 = x_1^{x_2^{x_3 * y}} \quad . \tag{16.1}$$

In this structure, x_3 plays the role of the model of the self. A statement that the subject presented by (16.1) has become aware of the self means that x_3 is substituted by the whole right-hand part of (16.1):

$$X_1 = x_1^{x_2^{x_1^{x_2^{x_3 * y}} * y}} \quad . \tag{16.2}$$

We will proceed to more general analysis. Consider the subject's readiness to choose the positive pole as a function of intention x_3:

$$F(x_3) = X_1. \tag{16.3}$$

The following sequence corresponds to the series of the acts of awareness:

$$\begin{array}{l} F(x_3) = X_1 \\ F(F(x_3)) = X_1 \\ F(F(F(x_3)) = X_1 \\ \dots\dots \end{array} \tag{16.4}$$

We will designate it

$$F^{(k)}(x_3) = X_1, \tag{16.5}$$

where $k = 1, 2, 3, ...$

Now we can generalize the concept of the Realist. The subject who chooses a value of x_3 for which the following equation holds

$$F^{(k)}(x) = x \tag{16.6}$$

will be called k-Realist.

We assume that the subject's cognitive system uses the iterative process

$$F^{(k)}(x_{(n)}) = x_{(n+1)} \tag{16.7}$$

for generating x.

We will show further that there are initial conditions under which the sequence (16.7) does not converge to a point, but the sequence

$$F^{(k+1)}(x_{(n)}) = x_{(n+1)} \tag{16.8}$$

does. In this way we will demonstrate that the subject's act of awareness makes him capable of taking a realistic decision.

Let us consider a few examples. Let the subject schematize the situation with function $M_3 = 1 - x$ and let $x_1 = 0$ and $x_2 = 0$. This subject corresponds to the sequence

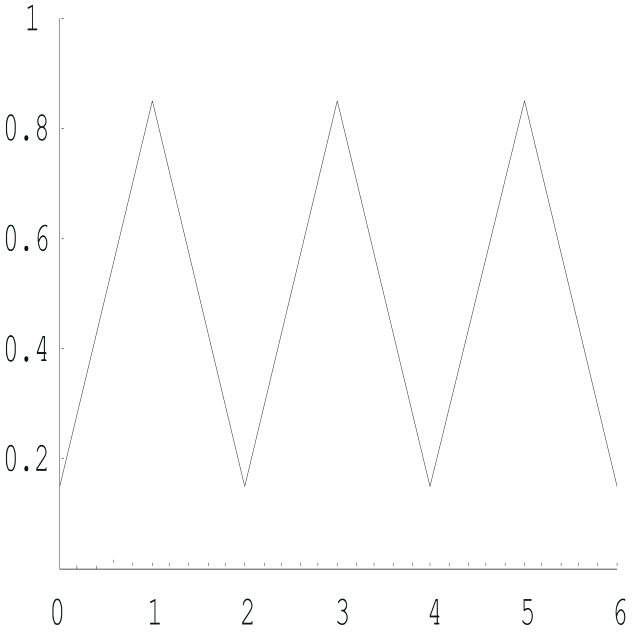

Figure 8: A process of oscillations generated by equation $x_{(n+1)} = 1 - x_{(n)}$, $x_{(0)} \neq (1/2)$.

$$\begin{array}{l} \qquad\quad 1 - x_{(n)} \\ \qquad\ \ 0 \\ F(x_n) = 0 \qquad\quad = x_{(n+1)}, \end{array} \tag{16.9}$$

or, in standard form,

$$1 - x_{(n)} = x_{(n+1)}. \tag{16.10}$$

The fixed point for (16.10) is $x = (1/2)$, which can be achieved only if $x_{(0)} = (1/2)$. With any other value, there begins an infinite process of oscillations of the subject's readiness as shown in Fig.8.

If this subject has completed an act of awareness, he will correspond to the following equation:

$$\begin{array}{l} \qquad\quad 1 - x_{(n)} \\ \qquad\ \ 0 \\ \quad 1 - 0 \\ \ \ 0 \\ 0 \qquad\quad = x_{(n+1)}, \end{array} \tag{16.11}$$

or, in standard form,

$$x_{(n)} = x_{(n+1)}. \tag{16.12}$$

Sequence (16.12) converges with any initial value of $x_{(0)} \in [0, 1]$. We see that one act of awareness has not only freed the subject corresponding to equation (16.9) of hesitations but, in addition, it has brought him to a state in which he is capable of free choice.

Consider now the subject corresponding to equation (15.7). This equation has a stable fixed point only at $a < 3$. When $a \geq 3$, bifurcation begins and a stable cycle of length 2 appears; that is, there are two special points and the members of sequence (15.8) tend to one of them when n is even and to the other when n is odd. Therefore there are oscillations of the subject's readiness. What happens when the subject performs an act of awareness? Let us write equation (15.8) as

$$F(x_{(n)}) = x_{(n+1)}. \tag{16.13}$$

Then, after the subject's act of awareness,

$$F(F(x_{(n)})) = x_{(n+1)}. \tag{16.14}$$

Note that sequences (16.13) and (16.14) are different. We use there the same expression $x_{(n)}$ only to avoid bulky notations. For any value of a from the interval $(3, 1+\sqrt{6})$, we can find two numbers c_1 and c_2 corresponding to the stable cycle. We will clarify now the difference between (16.13) and (16.14). Let us choose an arbitrary $x_0 = u_0$ from the open interval (0,1), where $x_0 \neq ((a-1)/a)$. Designate sequence (16.13) as $u_0, u_1, u_2, u_3, \ldots$. When n increases, the members of this sequence with even numbers, $u_0, u_2, u_4 \ldots$, converge to one of the two numbers, c_1 or c_2, and those with the odd numbers, $u_1, u_3, u_5 \ldots$, to the other. The sequence (16.14) consists of the members of sequence (16.13) having even numbers and, therefore, it has limit. Thus the act of awareness has stopped the subject's hesitations by generating a new stable fixed point. When a reaches $1+\sqrt{6}$, two bifurcations take place, the cycle of length 2 loses its stability, and a new cycle with length 4 appears. In this case one act of awareness does not save the subject from oscillations; he has to make two more consecutive acts of awareness, and as a result, sequence (16.14) turns into $F^{(4)}(x_{(n)}) = x_{(n+1)}$. The subject, by making consecutive acts of awareness according to the law

$$F^{(2^p)}(x_{(n)}) = x_{(n+1)}, \tag{16.15}$$

where $p = 1, 2, ...$, eliminates the oscillations of readiness at all values of $a < 3.57$. Beginning from this point the behavior of the sequence becomes so complicated that any "meditation" consisting of a finite number of the acts of awareness does not allow generation of a sequence converging to a stable fixed point.

We see here that one of the possible deep functions of awareness consists of a fight with chaos which may appear in consecutive cognitive computations. A subject who is not capable of being aware of the self loses stability at $a = 3$. A subject who can perform multiple acts of awareness keeps stability in the area of chaos until $a = 3.57$.

17 The Moment of Choice

converging streams of consciousness; awareness and the moment of choice

Let us make the concept of the subject's readiness more precise. Readiness is the probability of choice of the positive pole under the condition that the act of choice has taken place. If the act of choice has not yet taken place, however, readiness is a characteristic of the subject's uncertain state. A natural question arises: at what moment does an uncertain state collapse into the choice of a specific alternative? We can assume that the choice happens when the difference between $x_{(n)}$ and $x_{(n+1)}$ just drops below a fixed constant ϵ:

$$|x_{(n)} - x_{(n+1)}| < \epsilon. \tag{17.1}$$

The value $n+1$ will be called the *moment of choice* and designated $n*$.

Consider an example. The subject's readiness is given by the following equation:

$$x_1^{x_2^{1-x_{(n)}}} = x_{(n+1)}, \tag{17.2}$$

or, in standard form:

$$x_1 + (1-x_1)(1-x_2)(1-x_{(n)}) = x_{(n+1)}.$$

Let $x_1 = (1/2)$, $x_2 = (1/2)$, $x_{(0)} = (5/8)$, and $\epsilon = 0.001$. The graph of $x_{(n)}$ evolution is given in Fig.9.

By direct computations we find that equation (17.1) holds for the first time at the fourth iteration, therefore $n* = 4$.

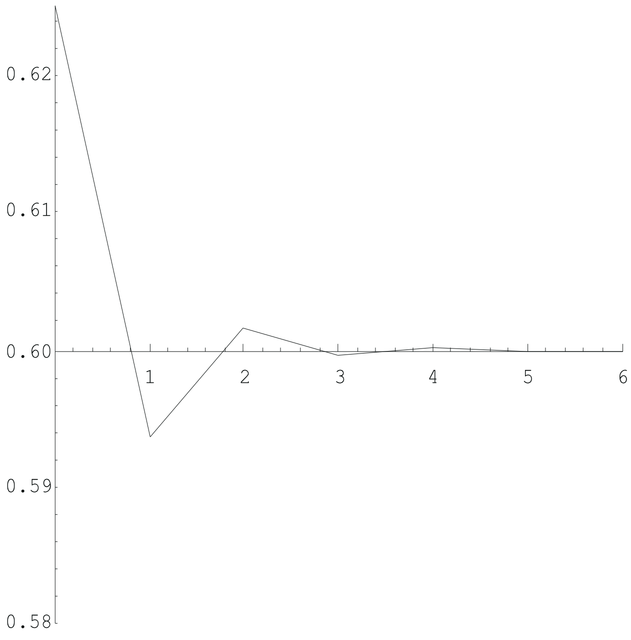

Figure 9: Damping of oscillations with increasing n. Choice is performed at the moment $n* = 4$.

Assume now that $x_1 = 0$ and $x_2 = 0$ and that the values of $x_{(0)}$ and ϵ are the same as they were previously. The subject corresponds to equation (16.9) and his stream of consciousness is in oscillation (see Fig.8). The values of $x_{(n)}$ swing between (5/8) and (3/8). The subject is incapable of making a choice. As was shown earlier, when the subject performs one act of awareness (see (16.11)) his stream of consciousness stops oscillating and the subject can make a choice. Awareness can also destroy a converging stream of consciousness and turn it into oscillation. Let the subject under consideration perform the second act of awareness. Then his stream of consciousness begins oscillating and he loses the ability to make a choice. Let us now find, which factors influence the moment of choice $n*$ if the stream of consciousness converges. It is clear that for fixed ϵ, $n*$ may depend on x_1, x_2, and x_0. It turns out that $n*$ can be determined completely by the number of acts of awareness, k. Let the following sequence converge at $x_{(0)}$:

$$F(x_{(n)}) = x_{(n+1)}. \tag{17.3}$$

For this value it is always possible to find k such that the sequence

$$F^{(k)}(x_{(n)}) = x_{(n+1)} \tag{17.4}$$

satisfies (17.1) for $n* = 2$. This means that the subject with a converging stream of consciousness is always capable of making a decision at moment 2. Note that this conclusion is based on the assumption that inner cognitive computations are done instantly.

18 Completion of the Classical Theory of Zero-Sum 2x2 Games

deontological aspect of games; Boolean matrix; Boolean functions; maximin mixed strategies as normalized utility-measures; model application in games

Classical game theory is based on a special schematization of the utilitarian aspect of choice; it does not take into account the deontological aspect. In this section we will try to supplement game theory by adding the deontological aspect. We will do so by using a static version of the theory described in the previous sections.

Table 23: 2x2 Zero-Sum Game Matrix

5	1
1	5

Consider the matrix of a zero-sum game shown in table 23. The rows correspond to X's strategies, and the columns to Y's. The numbers in the cells show X's payoffs; Y's payoffs are equal to X's payoffs taken with the opposite sign. Classical game theory advises each player to use maximin mixed strategy $(1/2, 1/2)$ in this game.

The game matrix represents only the utilitarian aspect of the game. We will now show how to complement the utilitarian aspect with the deontological aspect by introducing a special Boolean matrix.

Let the first strategy of each player be"to tell the truth" and the second"to lie." For each player the first strategy plays the role of the positive pole (positive strategy), the second that of the negative pole (negative strategy). Assume also that subject X considers his payoff of one unit too small and that subject Y considers his loss of five units too great. Thus, X assesses positively only those outcomes when he wins 5 units, and Y only those when he loses 1 unit.

Everything said above can be represented in a Boolean matrix (see Table 24).

The rows and columns of this payoff matrix correspond to the initial matrix. The Boolean values of 1 on the left and at the top corre-

Table 24: Boolean Matrix

	1	0
1	1,0	0,1
0	0,1	1,0

spond to the positive pole, 0 to the negative, and each pair of Boolean values in the matrix cells shows the subjects' assessment of the result corresponding to a given pair of strategies. The left value is X's assessment, the right one Y's assessment.

Using this matrix we can reconstruct the Boolean functions M_X and M_Y corresponding to the players:

Table 25: Boolean M-Functions for X and Y.

x_3	y	M_X	y_3	x	M_Y
1	1	1	1	1	0
1	0	0	1	0	1
0	1	0	0	1	1
0	0	1	0	0	0

It follows from the tables that $M_X = M_{15}$ and $M_Y = M_{14}$. If the players, X and Y, imitate each other exactly, the game corresponds to the following system of equations:

$$\begin{array}{ll} \quad M_{15}(x,y) & \\ x_2 & \\ x_1 & = x \\ & \\ \quad M_{14}(y,x) & \\ y_2 & \\ y_1 & = y. \end{array} \tag{18.1}$$

The values of x_1, x_2, y_1, and y_2 reflect the utilitarian aspect of the game, and we can extract them from the payoff matrix.

The main idea is that we introduce a new interpretation of maximin mixed strategies. We consider them to be pairs of normalized utility measures corresponding to the alternatives instead of pairs of probabilities of choosing the pure strategies.

Note that in their mathematical nature, maximin mixed strategies are not connected with the notion of probability. They are sets of non-negative numbers representing the solution of a linear problem of optimization and their sum is equal to 1. To consider those numbers as probabilities is only one possible interpretation. In the framework of our model, we use another interpretation: each number signifies pressure toward an alternative.

In the game which we are analyzing, maximin mixed strategies are $(1/2, 1/2)$, so we assume that $x_1 = (1/2)$ and $y_1 = (1/2)$. Let the game matrix reflect the utility-measures for the outcomes both at the moment of choice and prior to it, so that the actual utility-values coincide with the expected ones. In this case $x_1 = x_2 = (1/2)$ and $y_1 = y_2 = (1/2)$. The system of equations (18.1) can be rewritten as follows:

$$\begin{array}{ll} \qquad 1-x-y+2xy & \\ \quad (1/2) & \\ (1/2) & = x \\ & \\ \qquad x+y-2xy & \\ \quad (1/2) & \\ (1/2) & = y, \end{array} \tag{18.2}$$

or, in canonical form:

$$\begin{array}{l} (1/2)xy - (5/4)x - (1/4)y + (3/4) = 0 \\ (1/2)xy - (1/4)x + (3/4)y - (1/2) = 0. \end{array} \tag{18.3}$$

The canonical forms for X were given in section 9; to obtain them for Y, we have to reverse the places of x and y. The only solution of this system in the area $0 \le x, y \le 1$ is

$$x \approx 0.633, \quad y \approx 0.617.$$

Thus, according to our model, X and Y will use the mixed strategies

$$(0.633, 0.367), \quad (0.617, 0.383),$$

instead of the maximin mixed strategy $(0.5, 0.5)$ as predicted by the classical game theory. The general plan for using our model, when the matrix of a zero-sum game is known is as follows:

1. The conditions of interaction are analyzed to determine which pure strategies play the role of positive and negative poles for each subject and how they assess each of the four outcomes; Boolean 1 is assigned to positive assessments, and Boolean 0 to negative assessments.

2. Based on these data a Boolean matrix is constructed; functions $M_X(x, y)$ and $M_Y(x, y)$ are found.

3. Maximin mixed strategies $(p, 1-p)$ and $(q, 1-q)$ are computed in standard way; p corresponds to X's pure positive strategy and q to Y's pure positive strategy. It is assumed that $x_1 = x_2 = p$ and $y_1 = y_2 = q$.

4. A system of equations is constructed:

$$\begin{array}{l} \quad M_X(x,y) \\ \;p \\ p \quad\;\; = x \\ \\ \quad M_Y(y,x) \\ \;q \\ q \quad\;\; = y, \end{array} \tag{18.4}$$

and its solutions are found.

5. The solutions are interpreted in the context of the subjects' interaction.

The scheme given above has been constructed under the assumption that no surprises occur: a player plays the game he expected to play; that is, the payoff matrix does not change at the moment of choice. If a new matrix appears, computations of x_1 and y_1 are conducted based on it, but x_2 and y_2 remain computed with the"old" matrix.

19 Reflexive Games with Saddle Point

classical game theory predictions vs. reflexive game theory predictions

In classical 2x2 games with a saddle point, each player chooses a strategy corresponding to the saddle point with probability 1. Thus, in a corresponding reflexive game, p and q take on either 1 or 0. The value of 1 is taken when the saddle point is tied up with the positive strategy, and the value of 0 when it is tied up with the negative strategy. The following identity holds:

$$1^R \equiv 1, \tag{19.1}$$

where R is an arbitrary number from [0,1]. Therefore, if a saddle point corresponds to the positive strategy, the subject chooses a strategy having a saddle point with probability 1; that is, he behaves in the same way as a player in classical game theory. But if the subject's positive strategy does not corresponds to the saddle point, his behavior may differ significantly from that predicted by classical game theory. Consider the following matrix:

Table 26: Zero-Sum Game Matrix

5	2
2	1

The saddle point appears in X's 1st strategy and in Y's 2nd strategy. Let this matrix correspond to the following Boolean matrix:

Table 27: Boolean Matrix

	1	0
0	1,1	0,1
1	1,0	0,0

We find that $M_X = M_4$, $M_Y = M_5$. For each subject the saddle point corresponds to the negative strategy, so $p = 0$, $q = 0$, and the following system describes the game:

$$\begin{matrix} y \\ 0 \\ 0 \end{matrix} = x \qquad\qquad\qquad \begin{matrix} 1-x \\ 0 \\ 0 \end{matrix} = y. \tag{19.2}$$

The standard form of this system is

$$\begin{aligned} x - y &= 0 \\ x + y &= 1. \end{aligned} \tag{19.3}$$

Thus $x = (1/2)$ and $y = (1/2)$, i.e., each player chooses the strategy corresponding to the saddle point with probability 0.5. This example shows that, in the theory of reflexive games, the probability of choosing the strategy with the saddle point can differ from 1.

It is possible to construct a game having a saddle point in which players have freedom of choice. Let the payoff matrix given above correspond to the following Boolean matrix:

Table 28: Boolean Matrix

	1	0
0	0,1	0,0
1	1,1	1,0

We find that $M_X = M_2$, $M_Y = M_2$ and obtain

$$\begin{matrix} x \\ 0 \\ 0 \end{matrix} = x \qquad\qquad\qquad \begin{matrix} y \\ 0 \\ 0 \end{matrix} = y. \tag{19.4}$$

It follows from (19.4) that x and y can take on any values from the interval [0,1].

20 Risk and Caution

concept of danger; inclination to risk; inclination to caution; Generous Maximalist; stingy player

Classical game theory describes the behavior of a cautious player. The problem of the formal description of conditions under which players are ready to risk has not been solved in the framework of this theory. In this section we will show a possible approach to this problem with the help of the model of the subject with reflexion.

In our everyday experience, the concept of risk is connected with the concept of danger. To risk means to make a choice which increases the chance of loosing something. Not every situation has a potential danger. So we will start by introducing a formal analogue of danger into the theory of 2x2 zero-sum games. A potential payoff is considered maximal for a player if it is strictly higher than any other payoff. Let us introduce the following definitions. A strategy is called *dangerous* if, in the framework of classical game theory, it must be chosen with a probability less than (1/2), but only players choosing it have a chance of receiving the maximal payoff. Consider for example the following matrices:

Table 29: *Danger and Non-Danger* Matrices. In the game with matrix (a) strategy 1 is dangerous for X; in the game with matrix (b) there is no dangerous strategy for X.

4	-5	1/10	4	-10	4/11
1	2	9/10	-2	6	7/11
	(a)			(b)	

The probabilities of choosing each pure strategy by player X according to classical theory are shown on the right hand sides. For X, the maximal potential payoff in matrix (a) is equal to 4 units; it may be reached only if X chooses strategy 1. This strategy, in accordance with the classical game theory, must be chosen with a probability of (1/10), that is, lower than (1/2), so, by our definition, it is dangerous for X. The danger is X's loss of 5 units if Y chooses strategy 2. In the game with matrix (b), there is no dangerous strategy for X, because his maximal payoff of 6 units can be reached only with strategy 2, which must be chosen with a probability of (7/11), that is, greater than (1/2).

We say that a player is *inclined* to risk if he is ready to choose a dangerous strategy with a probability greater than (3/4). The reason of choosing this constant value will be explained later.

We say that a player is *cautious* if he chooses a strategy in accordance with classical game theory.

Further, we will show that in the framework of reflexive game theory, both kinds of players exist: players inclined to risk and players inclined to caution. Let us introduce some general characteristics of players. We call a player a *Maximalist* if he always considers the strategy with a potentially maximal payoff to be the positive pole. For example, in the game with matrix (a), the positive pole for the Maximalist is strategy 1, and in matrix (b) it is strategy 2. It is clear that a dangerous strategy always plays the role of the positive pole for the Maximalist.

Let us call a player *generous* if he assesses any outcome positively, i.e., he schematizes situations with function $M_{16} \equiv 1$; let us call a player *stingy* if he assesses any outcome negatively, i.e., he schematizes situations with function $M_1 \equiv 0$.

To explain the motive for our choice of the terminology, let us consider matrix (a). We begin with player X, who assesses all the outcomes positively. Let him choose strategy 1. If Y selects strategy 1, X receives 4 units, so it is natural for him to consider this result positive. If Y selects 2, X looses 5 units; if he also assesses this result positively, this means that he easily (at least in his mind) parts with his resources, so the term *generous* is appropriate to characterize him.

Consider now another X, who assesses all the outcomes negatively. Let him also choose strategy 1 in matrix (a). If Y chooses strategy 2, X must give him 5 units; it is clear that he assesses this result negatively. If Y chooses strategy 1, X receives 4 units; if he assesses this result also negatively, we can suppose that he consider this payoff too small, so we call him stingy.

Now we will demonstrate that in games with dangerous strategy without surprise the Generous Maximalist is always inclined to risk. The absence of surprise means that $x_1 = x_2$. Then the following equation corresponds to the Generous Maximalist in the game with dangerous strategy:

$$\begin{array}{l} \quad M_{16}(x,y) \\ \quad p \\ p \qquad = x, \end{array} \qquad (20.1)$$

where $p < (1/2)$. Equation (20.1) can be rewritten as

$$\begin{array}{l} \quad 1 \\ \quad p \\ p \qquad = x, \end{array} \qquad (20.2)$$

or, in standard form, as

$$x = 1 - p + p^2. \qquad (20.3)$$

Function (20.3) reaches its minimum at $p = 1/2$, and its minimum is equal to (3/4). For $p < (1/2)$, the value of $x > (3/4)$ and, therefore, the Generous Maximalist is always inclined to risk. For example, in the game with matrix (a), such a player would choose strategy 1, which is dangerous, with a probability of

$$x = 1 - (1/10) + (1/10)^2 = 0.91. \qquad (20.4)$$

We see that the constant of (3/4) introduced in our definition of inclination to risk is the minimum of function (20.3).

Let us proceed to a discussion of caution. We will demonstrate that a stingy player chooses strategies with those exact probabilities which are prescribed by classical game theory. The following equation corresponds to a stingy player:

$$\begin{matrix} & M_1(x,y) & \\ x_2 & & \\ x_1 & & = x, \end{matrix} \tag{20.5}$$

or

$$\begin{matrix} & 0 & \\ x_2 & & \\ x_1 & & = x, \end{matrix} \tag{20.6}$$

whence

$$x = x_1, \tag{20.7}$$

where x_1 is the probability with which, in the framework of classical game theory, the alternative playing the role of the positive pole must be chosen. Thus, stingy player is inclined to be cautious.

21 Generalization to Non-Zero-Sum Games

maximin as a basis for computing x_1 and x_2, general scheme, prisoner's dilemma

In section 18 we established connections between the formal model of the subject and zero-sum games. In this section, we will expand these connections to non-zero-sum games 2x2. We will begin by elaborating the problem of how X's cognitive system takes into consideration the presence in the situation of another subject, Y. Within the framework of our representation, the presence of another subject is expressed in two ways: by using the model of the situation $M(x, y)$ and/or through imitation of the other subject's decisions (see section 13).

The values of x_1 and x_2 reflect the properties of concrete alternatives and do not depend on the presence or absence of another subject in the situation. So let us assume that the subject's cognitive system computes only its own maximin mixed strategies in finding the values of x_1 and x_2 without analyzing whether or not equilibrium is accessible within a given game. This assumption allows us to connect reflexive games with arbitrary classical games using matrices of the type:

Table 30: Payoff Matrix

a_1, b_1	a_2, b_2
a_3, b_3	a_4, b_4

The real numbers a_i are X's payoffs, and the real numbers b_i are Y's payoffs.

A general scheme for modeling the decisions of subjects in games not involving surprise and with an arbitrary payoff matrix 2x2 is as follows:

1. A Boolean matrix (Table 31) is constructed based on the analysis of the interaction of particular subjects (see section 18).

Table 31: Boolean Matrix

	b	$\overline{b}$
a	α_1, β_1	α_2, β_2
$\overline{a}$	α_3, β_3	α_4, β_4

Variables a, b, α_i, β_i take on Boolean 1 or 0 (in general independently of each other). Using this matrix we find the functions $M_X(x, y)$ and $M_Y(x, y)$ (see section 18). It is easy to see that for each payoff matrix there are $2^{10} = 1024$ possible different Boolean matrices.

2. The matrix in Table 30 is represented as two separate payoff matrices (see Table 32).

The first matrix allows us to find X's maximin mixed strategy $(p, 1 - p)$, and the second matrix gives Y's maximin mixed strategy $(q, 1 - q)$. The numbers p and q correspond

Table 32: Matrices for X and Y

a_1	a_2	b_1	b_2
a_3	a_4	b_3	b_4

to the subjects' positive pure strategies. It is assumed that $x_1 = x_2 = p$ and $y_1 = y_2 = q$.

3. The system of equation (18.4) is constructed and its solutions (x, y) are found.

For games involving surprise we have to establish two matrices of the type given in Table 30. One matrix reflects the preliminary payoff estimates and allows the computation of x_2 and y_2. The second matrix reflects the payoff information received by the subjects at the moment of choice; on the basis of this matrix the values of x_1 and y_1 can be computed. As an example let us analyze a version of the famous prisoner's dilemma. Two suspects are arrested. Each suspect has the choice of pleading either guilty or not guilty. If both of them plead guilty, they will both be sentenced to 3 years in prison; if both plead not guilty, they will serve 2 years; if one of them pleads guilty and the other not guilty, the first one will be sentenced to 1 year and the second to 50 years. These conditions are reflected in the payoff matrix given in Table 33 (each subject's first strategy corresponds to a not guilty plea and the second one to a guilty plea).

Table 33: Prisonner's Dilemma

-2,-2	-50,-1
-1,-50	-3,-3

Suppose now that these people have not committed any crime and were arrested by mistake. In this case, the alternative "to plead guilty" is a lie and as such figures as the negative pole; the alternative "not to plead guilty" is the truth figures as the positive pole. We assume that the prospect of spending 50 years in prison is perceived as a catastrophe by each subject, so that in comparison the terms of 1, 2 or even 3 years seem promising outcomes. Thus the following Boolean matrix corresponds to this game:

Table 34: Boolean Matrix

	1	0
1	1,1	0,1
0	1,0	1,1

Using this matrix we find that $M_X = M_{12}$ and $M_Y = M_{12}$; thus

$$M_X = 1 - x + xy$$
$$M_Y = 1 - y + xy. \tag{21.1}$$

Let us represent the payoff matrix (Table 33) as two matrices (Table 35):

Table 35: Matrices for X and Y

-2	-50	-2	-1
-1	-3	-50	-3

Using these matrices we find that X's and Y's maximin mixed strategies are $(p = 0, \quad 1 - p = 1)$ and $(q = 0, \quad 1 - q = 1)$. As a result we obtain the following system of equations:

$$\begin{array}{l} \quad 1 - x + xy \\ \;0 \\ 0 \qquad\quad = x \\ \\ \quad 1 - y + xy \\ \;0 \\ 0 \qquad\quad = y, \end{array} \tag{21.2}$$

or, in standard form:

$$1 - 2x + xy = 0$$
$$1 - 2y + xy = 0. \tag{21.3}$$

The system (21.3) has only one solution ($x = 1, y = 1$). Therefore, the model predicts that the suspects will decide to plead not guilty so that each of them will be sentenced to 2 years of prison time.

We have to emphasize that this conclusion essentially depends on the Boolean matrix (Table 34), which in turn reflects the deontological aspect of the problem. If another Boolean matrix were used the prediction might be different.

22 Model of the Reflexive Subject and Experimental Psychology

Adams-Webber's experiment; the golden section ratio; mere- exposure experiments; psychophysical experiments

The model of the reflexive subject is based on a few very general assumptions about human choice. To what extent do the model's predictions correspond to reality? In this section we describe the results of experimental testing of some of the model's predictions and demonstrate that it can explain certain phenomena which have remained unclear for a long time.

In the 1970s Adams-Webber and Benjafeld (1973) discovered that people choose positive poles with the frequency 0.62 while evaluating their friends and acquaintances on bipolar constructs (of the type strong-weak). They also proposed the hypothesis that the exact value of this number is equal to the golden section ratio (Benjafeld & Adams-Webber, 1976). In the 1980s the present author noted the fact that the value 0.62 appears only when the subjects choice cannot be based on clear operational criteria. This observation allows us to propose the following explanation of the phenomenon of the"golden section" (Lefebvre, 1985; 1987).

When there are no operational criteria for choice, the subject's cognitive system is unable to compute the values of pressure toward the positive and the negative poles between which he is to choose. Thus x_1 takes on the neutral value of (1/2). Next there begins the iterative process of the computation of the subject's readiness to choose the positive pole:

$$(1/2)^{x_{(n)}^{x_{(n)}}} = x_{(n+1)}, \tag{22.1}$$

where $x_{(0)}$ is an arbitrary number from $[0, 1]$.

Equation (22.1), in standard form, is as follows:

$$\begin{aligned} x_{(n+1)} &= -(1/2)x_{(n)}^2 \\ &+(1/2)x_{(n)} + (1/2). \end{aligned} \tag{22.2}$$

This sequence converges at $n \to \infty$ (see, for example, Lefebvre, 1982, p.139), and its limit x is a positive root of the equation $x^2+x-1 = 0$, i.e.

$$x = \frac{\sqrt{5}-1}{2} = 0.618... \tag{22.3}$$

The golden section results from the iterative process (22.1). In accordance with the model, such a process must appear when two conditions are satisfied: (a) alternatives are polarized, that is, one alternative plays the role of the positive pole and the other one that of the negative pole; and (b) there is no criterion for the utilitarian preference of one alternative over the other. These conditions are present in experiments with the"mere exposure" phenomenon (Zajonc, 1968). The essence of this phenomenon is that the subjects, when confronted with a series of practically equivalent alternatives, choose more frequent and evaluate more highly those which were presented earlier. If we suppose that the preliminary exposure to an alternative turns it into the positive pole, then in accordance with our model it must be choh

Table 36: Experimental data exhibiting the"mere exposure" phenomenon.

Kunst, Wilson, et al. (1980)	0.60
Seamon, et al. (1983)	0.61
Mandler, et al. (1987)	0.62
Bonanno, et al. (1986)	0.66
Bonanno, et al. (1986)	0.63
Bonanno, et al. (1986)	0.62
Bonanno, et al. (1986)	0.61
Bonanno, et al. (1986)	0.63
Bonanno, et al. (1986)	0.62

sen with the frequency 0.62. We have analyzed all available work on the"mere exposure" phenomenon (Lefebvre, 1995; 1997a). Table 36 shows the experimental frequencies of the choice of alternatives to which the subjects were exposed in advance; each figure corresponds to a separate experiment. We see that all the values are around 0.62.

Equation (22.1) allows us to obtain more detailed predictions about frequencies appearing in bipolar choice. Let us assume that the subject has *intention* to evaluate the alternative playing the role of the positive pole as strictly positive; that is $x_3 = 1$. Then equation (22.1) changes into

$$\begin{matrix} 1 \\ x_{(n)} \\ (1/2) = \quad x_{(n+1)}, \end{matrix} \tag{22.4}$$

where $x_{(0)} \in [0,1]$. Equation (22.4), in standard form, is as follows:

$$x_{(n+1)} = 1 - (1/2)x_{(n)}, \tag{22.5}$$

whence

$$x_{(n)} \to (2/3). \tag{22.6}$$

Now let the subject have *intention* to evaluate the alternative playing the role of the positive pole as strictly negative, that is, $x_3 = 0$. Then equation (22.1) turns into the identity

$$x_{(n)} \equiv (1/2), \tag{22.7}$$

that is, in this case the subject will choose the poles with equal probability. Adams-Webber (1997) tested these predictions experimentally (see also Adams-Webber & Rodney, 1983). The subjects in his experiments evaluated not only their acquaintances but also themselves using the same bipolar constructs. If we suppose that, for each pair of adjectives, the subjects have intention to evaluate others in the same way as themselves, then, given a positive self-evaluation, the frequency of positive evaluations of others is equal to (2/3), as follows from (22.6), and given a negative self-evaluation, the frequency of positive evaluations of others is equal to 0.5, as follows from (22.7). Adams-Webber has found that the frequencies of positive evaluations of others are as follows:

for positive self - 0.67; 0.67; 0.68

for negative self - 0.5; 0.48; 0.52

As we see, the reflexive model passes this test.

The value of readiness measures the subject's uncertain state. We presume that this value can reveal itself not only in the frequency of the subject's positive choices but also in the subject's judgments (Lefebvre, 1987). Let the subject be ready to call a given stimulus"strong" with the probability x, but the experimenter asks him not to make an actual choice but to mark a point on a scale [0,1] showing to what degree a given stimulus can be considered strong. We assume that, in such cases, the subject "projects" his readiness onto the scale [0,1] and indicates point x. A broader interpretation of the value of the subject's readiness allows us to use the model not only for investigations of choice, but also for studies of the processes of estimation.

Based on the broadened model we wish to propose an explanation for the paradoxical phenomenon discovered by Poulton et al. (1968). In these experiments, the subjects were presented with three samples of paper: black, white, and gray (the"lightness" of the gray paper midway between the other two on the psychological scale). The subjects had to estimate the degree of lightness of the gray sample on a scale of 0 to 100 (0 being black, 100 being white); each subject performed the estimation only once, and only the first touch of a pencil to the paper was registered.

It would seem that the distribution of the subjects' estimations should look roughly as in Fig.10a. But the experiment resulted in the graph shown in Fig.10b. Instead of the expected peak at the point 50, there is a hole. This paradoxical result has not been explained within the framework of traditional psychophysics. The essence of our explanation is as follows. Each subject has only two samples, black and white, on which to base his estimation. This information is insufficient for the subject's cognitive system to construct an individual scale of lightness and compute the value of the pressure toward the positive pole. Thus the iterative process (22.1) begins and generates a value of readiness equal to the golden section. The conditions of the experiment did not determine a correspondence of the black and white samples to positive and negative poles. Therefore, for some subjects, the white sample played the role of the positive pole, and for others, the black one. The former marked 0.62 of the scale from the point 0, and the latter from the point 100. The resulting graph is a summary of two distributions: one with the peak at 0.62, and the other at 0.38 (Fig.11) (see Lefebvre, 1992b).

Consider now another experiment of the category rating of stimuli. The subjects are presented first with the strongest and the weakest stimuli from a set and then with all others one by one. The task is to relate each stimulus to a category, for example, the weakest to category 1, the strongest to category 11, and all others to different intermediate categories. Stevens & Galanter (1957) found that the connection between category rating and direct magnitude estimation is not linear (Fig.12a). Moreover, they discovered that when the weak stimuli are presented more often than the strong ones, the curve is more convex, and when the strong stimuli are presented more often than the weak ones, the

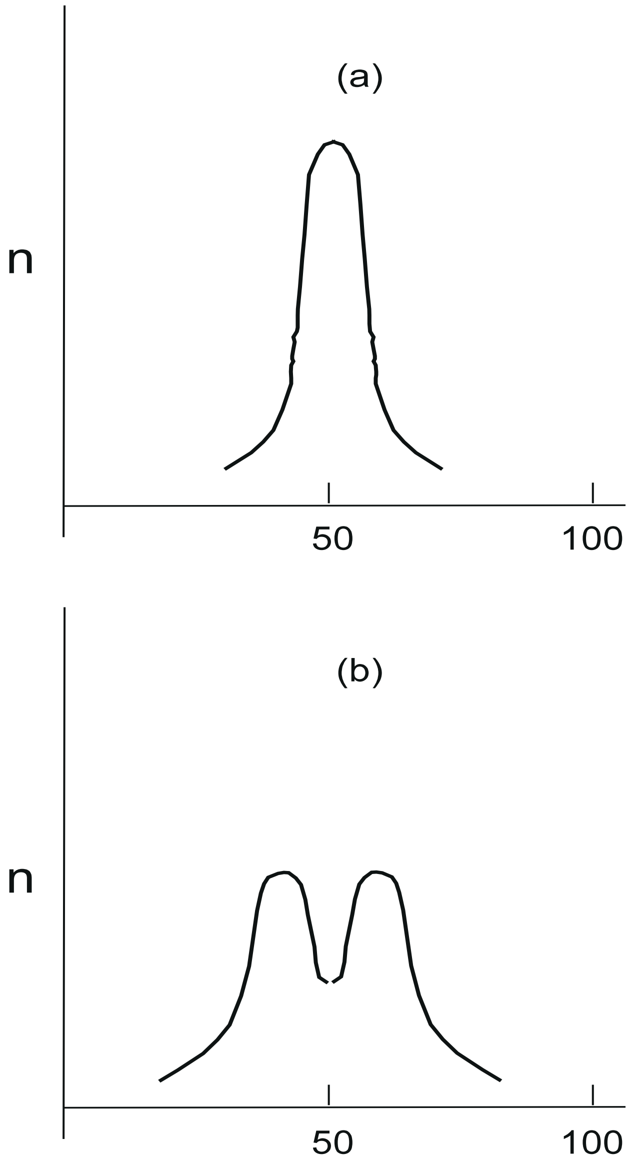

Figure 10: Distribution of estimations of the lightness of the gray sample. (a) Expected distribution; (b) actual experimental distribution.

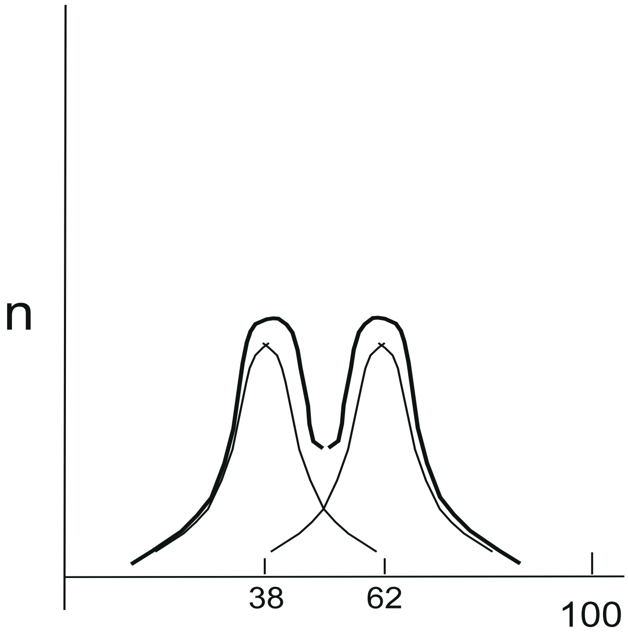

Figure 11: Explanation of double peaks of the distributions of individual estimations of lightness.

curve approaches a straight line (Fig.12b).

The reflexive model allows us to suggest an explanation for these phenomena (Lefebvre, 1992; 1997a). Consider the equation for a Realist:

$$x$$
$$x_2$$
$$x_1 \quad = x. \qquad (22.8)$$

Variable x represents a normalized category rating ($x = 0$ corresponds to category 1, and $x = 1$ to a category with the maximal number k). Variable x_1 shows a stimulus normalized magnitude estimation, and variable x_2 is the expected magnitude estimation at the moment of category rating. The value of x_2 can be interpreted as the mean value of normalized magnitude estimations of all the stimuli preceding a given one. For a long, well randomized series, x_2 can be considered a constant value. The solution for equation (22.8) is

$$x = \frac{x_1}{x_1 + x_2(1 - x_1)}. \qquad (22.9)$$

It is easy to see that equation (22.9) corresponds to the family of hyperbolas with parameter x_2 shown in Fig.13.

The more often weak stimuli are presented in a given series, the smaller the value of x_2 and the more convex the hyperbola which cor-

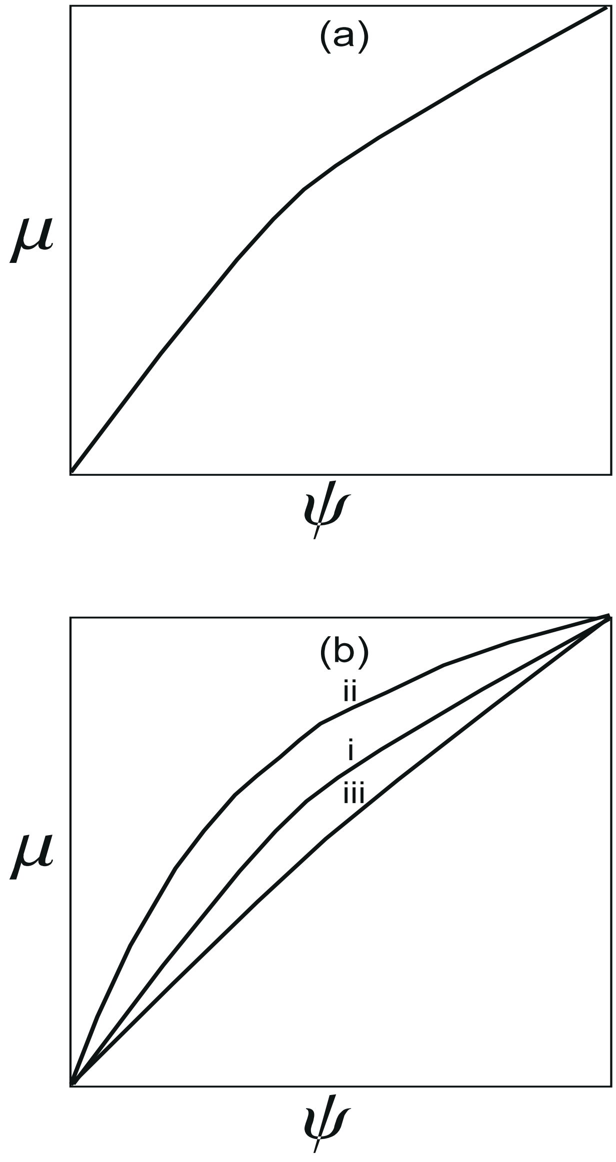

Figure 12: Connection of categorical (μ) and magnitude (ψ) estimations. (a) Equal distribution of weak and strong stimuli in the experimental series. (b) Line i represents an equal distribution, line ii shows a distribution shifted toward weak stimuli, and line iii corresponds to a distribution shifted toward strong stimuli.

responds to the experimental data. Fig.14 displays two experimental graphs of category rating and the theoretical hyperbolas corresponding to them.

Therefore, with the help of the reflexive model we have deduced hypothetical hyperbolic law of stimuli categorization.

All the experiments cited above support the claim that the model of the reflexive subject reveals some features of real human cognitive processes.

Acknowledgement

I would like to express my appreciation to Dr. Victorina Lefebvre and Dr. L. D. Miller for numerous productive discussions and for their invaluable help in preparation the manuscript.

23 References

Adams-Webber, J. (1987). Comment on Lefebvre's Model from the Perspective of Personal Construct Theory. *Journal of Social and Biological Structures*, **10**, 177-189.

-" - (1995). A Pragmatic constructivist Gambit for Cognitive Scientists. *PSYCOLOQUY* **6**(34).

-" - (1997). Self-reflexion in Evaluating Others. *American Journal of Psychology*, **110**, 527-541. Adams-Webber, J. & Ben-

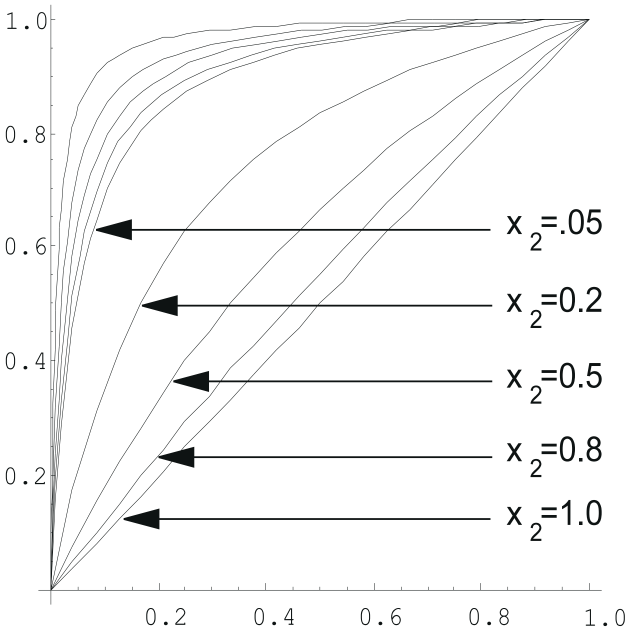

Figure 13: The family of hyperbolas. The smaller x_2, the more convex the curve.

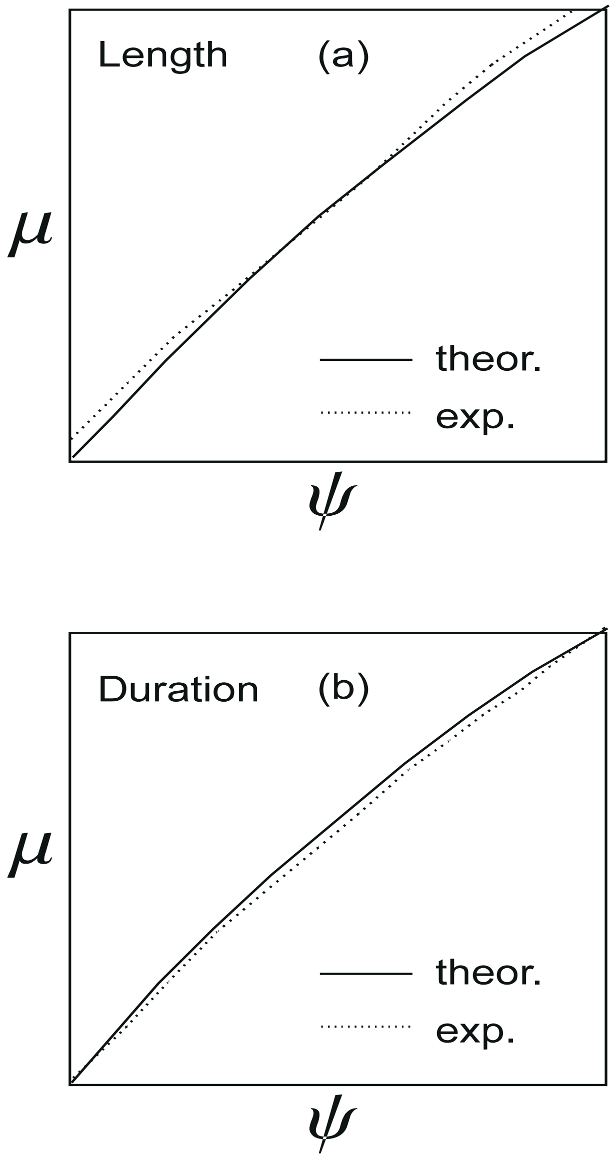

Figure 14: Theoretical and experimental graphs of category rating. (The experimental graphs were constructed according to the data of Stevens & Galanter, 1957.)

jafield, J. (1973). The Relation between Lexical Marking and Rating Extremity in Interpersonal Judgment. *Canadian Journal of Behavioral Sciences*, **5**, 234-241.

Adams-Webber, J. & Rodney, (1983). Rational Aspects of Temporary changes in Construing Self and Others. *Canadian Journal of Behavioral Sciences*, **15**, 52-59.

Batchelder, W. W. (1987). Comments on Some Critical Issues in Lefebvre's Framework for Ethical Cognition. *Journal of Social and Behavioral Structures*, **10**, 214-226.

Benjafield J. & Adams-Webber, J. (1976). The Golden Section Hypothesis. *British Journal of Psychology*, **67**, 11-15.

Bonnano, G. A. & Stilling, N. A. (1986). Preference, familiarity and recognition after repeated brief exposure to random geometric shapes. *American Journal of Psychology*, **99**, 403-415.

Fink, J. G. (1980). *Depersonalization and Personalization as Factors in a Taxonomy of Acting.* Dissertation, New York University, University Microfilms International.

Kauffman, L. (1990). Self and Mathematics. In Wheeler, H. (Ed.), *The Structures of Human Reflexion*, New York: Peter Lang.

Krylov, V. Yu. (1994). On One Model of Reflexive Behavior Distinct from Lefebvre's Model. *Applied Erconomics*, **1**, 21-24.

Kunst-Wilson, W. R. & Zajonc, R. B. (1980). Affective Discrimination of Stimuli that Cannot Be Recognized. *Science*, **207**, 557-558.

Lefebvre, V. A. (1965a). Basic Ideas of the Logic of Reflexive Games. In: *Problemy Issledovania Sistem i Struktur.* Moscow: Academy of Sciences of the USSR Press.

-" - (1965b). On Self-reflexive and Self-organizing Systems, In: *Problemy Issledovania Sistem i Struktur.* Moscow: USSR Academy of Sciences Press. English translation in *Applied Ergonimics* (1994), **1**, 6-9.

-" - (1966). Elements of Logic of Reflexive Games. In: *Problemy Inzhenernoi Psichologii.* Leningrad, **4**, 273-279.

-" - (1967). *Konfliktuyushchie Struktury* (Conflicting Structures). Moscow: Vysshaya Shkola. (1973). Moscow: Radio (new enlarged edition).

-" - (1972). A Formal Method of Investigating Reflexive Processes. *General Systems*, **XVII**, 181-188.

-" - (1977a). *The Structure of Awareness.* Beverly Hills: Sage.

-" - (1977b). Introreflective Analysis: Using Algebraic Polynomials with Boolean Coefficients to Represent a Person's Inner World. *Behavioral Science*, **22**, 49-52.

-" - (1977c). A Formal Approach to the Problem of Good and Evil. *General Systems*, **22**, 183-185.

-" - (1979). Mathematical Modeling of ethical Systems. *Proceedings of the Second International Conference on Mathematical Modeling*, **2**, 719-727.

-" - (1980). An Algebraic Model of Ethical Cognition. *Journal of Mathematical Psychology*, **22**, 83-120.

-" - (1982). *Algebra of Conscience.* Dordrecht, Holland: Reidel.

-" - (1984). Structures and Function of Natural Reflexion. *Proceedings of the International Conference.* Society for General Systems Research, **2**, 359-363.

-" - (1985). The Golden Section and an Algebraic Model of Ethical Cognition. *Journal of Mathematical Psychology*, **29**, 289-310.

-" - (1987). The Fundamental Structures of Human Reflexion. *Journal of Social and Behavioral Structures*, **10**, 129-175.

-" - (1990). The Fundamental Structure of Human Reflexion (revised). In: Wheeler, H. (Ed.), *The Structure of Human Reflexion: The Reflexional Psychology of Vladimir Lefebvre.* New York: Peter Lang.

-" - (1992a). *A Psychological Theory of Bipolarity and Reflexivity.* Lewinston: The Edwin Mellen Press.

-" - (1992b). A Rational Equation for Attractive Proportions. *Journal of Mathematical Psychology*, **36**, 100-128.

-" - (1994). From the Representation of Rationality to the Representation of Free Will. *Psichologihesky Journal*, **15**, 2, 99-121.

-" - (1995). The Anthropic Principle in Psychology and Human Choice. *PSYCOLOQUY* **6**(29).

-" - (1996). Extracting Information about Subjective States from a Function Describing Behavior. *PSYCOLOQUY* **7**(8).

-" - (1997a). *The Cosmic Subject.* Moscow: Russian Academy of Sciences Institute of Psychology Press.

-" - (1997b). The Principle of Freedom. *PSYCOLOQUY* **8**(9).

Lefe[b]vr[e], V. A., Baranov, P. V., & Lepskiy, V. Ye. (1969). Internal Value in Reflexive Games. *Engineering Cybernetics*, No.4, 27-31.

Levitin, L. (1987). Bohr's Atom and Vladimir Lefebvre's "Inner Computer." *Journal of Social and Biological Structures*, **10**, 183-185.

Mandler, G., Nakamura, Y., & VanZandt, B. J. S. (1987). Nonspecific Effects of exposure on stimuli that cannot be recognized. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **13**, 646-648.

May, R. H. (1976). Simple Mathematical Models with Very complicated Dynamics. *Nature*, **261**, 459-467.

McClain, E. (1987). Comment on Vladimir Lefebvre's Tonal Automata. *Journal of Social and Biological Structures*, **10**, 2, 204-212.

Mickens, R. e. (1990). *Difference Equation.* New York: Van Nostrand Reinhold.

Miller, L. D. (1997a). *Logic Transformation.* Manuscript.

-" - (1997b). *Psy-Ops Applications of Reflexive Behavior Models.* Manuscript.

-" - (1997c). *New Axioms for Lefebvre's Logic and Variants.* Manuscript.

Pope, K. S. and Singer, I. L. (Eds.) (1978). *The Stream of consciousness.* New York and London: Plenum Press.

Popper, K. (1992). Note in the Appendix to Lefebvre, V. A., *A Psychological Theory of Bipolarity and Reflexivity.* Lewinston: The Edwin Mellen Press.

Poulton, E. S. & Simmonds, D. C. V. (1985). Subjective Zeros, Subjectivity Equal Stimulus Spacing, and Contraction Biases in Very First Judgments of Lightness. *Perception & Psychophysics*, **37**, 420-428.

Rapoport, A. (1982). Foreword. In Lefebvre, V. A. (Author), *Algebra of Conscience: A Comparative Analysis of Western and Soviet Ethical Systems.* Boston: Reidel.

-" - (1990). Reflexion, Modeling, and Ethics. In Wheeler, H. (Ed.): *The Structure of Human Reflexion.* New York: Peter Lang.

-" - (1995). Human Reflexion and the Anthropic Principle. *PSYCOLOQUY* **6**(37).

-" - (1996). Probabilistic choice and Lefebvre's Model of Human Reflexion. *PSYCOLOQUY* **7**(27).

Schreider, J. A. (1994). Fuzzy Sets and the Structure of Human Reflexion. *Applied Ergonomics*, **1**, 19-20.

-" - (1998a). *Nepreryvnye Logiki kak Yazyki Refleksii.* Manuscript.

-" - (1998b). *Etika* (Ethics). Moscow: Tekst.

Seamon, J. G., Brody, N., & Kauff, D. M. (1983). Affective Discrimination of Stimuli that are not Recognized: Effects of Shadowing, Masking, and Cerebral Laterality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **9**, 544-555.

Stevens, S. S. & Galanter, E. H. (1957). Ratio Scales and Category Scales for a Dozen Perceptual continua. *Journal of Experimental Psychology*, **54**, 6, 377-411.

Taran, T. A. (1997). Description of Reflexive Behavior by Means of Argumentation Logic. *Informatsionnye Protsessy i Sistemy*, No.5, 18-23.

Townsend, J. T. (1983). Algebra of Conscience (book review). *Journal of Mathematical Psychology*, **27**, 461-471.

-" - (1990). Lefebvre's Human Reflexion and Its Scientific Acceptance in Psychology. In Wheeler, H. (Ed.), *The Structures of Hu-*

man Reflexion, New York: Peter Lang.

Wheeler, H. (1987). A Constructional Biology of Hermeneutics. *Journal of Social and Biological Structures*, **10**, 2, 103-123.

-" - (1990a). A Reflexional Model of the Mind's Hermeneutic Processes. In Wheeler, H. (Ed.): *The Structure of Human Reflexion.* New York: Peter Lang.

-" - (1990b). *The Structure of Human Reflexion: Reflexional Psychology of Vladimir Lefebvre.* New York: Peter Lang.

Zajonc, R. B. (1968). Attitudinal Effects of Mere Exposure. *Journal of Personality and Social Psychology. Monograph supplement.* **9**, 2, part.2, 1-32.

-" - (1987). Comment on Interpersonal Affiliation and the Golden Section. *Journal of Social and Biological Structures*, **10**, 212-214.