The Law of Self-Reflexion:
A Possible Unified Explanation for the Three Different Psychological Phenomena

Vladimir A. Lefebvre
University of California, Irvine
USA

ABSTRACT

The centuries-old philosophical idea that man has an image of the self containing an image of the self (of the second order) obtains a new life in the mathematical model of the subject possessing reflexion. One assumption underlying the model is that the subject tends to generate patterns of behavior such that some kind of similarity is established between the subject himself and his second order image of the self. We demonstrate that this model allows a single explanation for three diverse, experimentally observed phenomena: (a) the nonlinear relation between magnitude estimation and categorization of identical stimuli (Parducci, Stevens, Galanter), (b) the avoidance of the value of 0.5 in estimating stimuli equidistant from two samples on a psychological scale (Poulton, Simmonds), and (c) the formal correspondence between, on the one hand, frequency of choice for particular alternatives and, on the other, reinforcement rate, found in some experiments with animals and people (Herrnstein, Baum). The results obtained allow us to hypothesize that the reflexive metaphor represents a general principle for regulation of both human and animal behavior.

Beginning with John Locke, the human ability to represent mentally one's own thoughts and feelings has been a central topic of Western philosophy (1, 2). This ability is conventionally called reflexion. The subject possessing reflexion can be depicted as a miniature human figure with the image of the self inside his head (Fig.1).

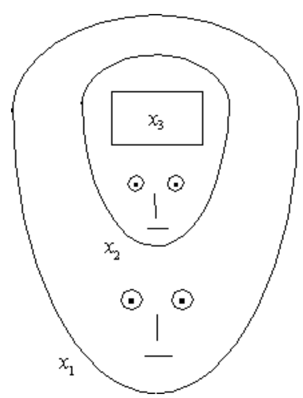


Fig. 1. The subject with reflexion. Inside the subject's inner domain, there is an "image of the self" with its own inner domain. An "image of the self" is traditionally regarded at as the result of the subject's conscious constructive activity. In the framework of the formal model of the subject with reflexion, the image of the self is not a product of human intellectual efforts, it is rather generated by the automatic work of the subject's cognitive mechanism (6, 13).

This image may contain thoughts and feelings including a description of the self; that is, this figure not only sees himself but also sees himself seeing himself. Although the idea of reflexion played an important role in nineteenth century psychology, it has not become part of mainstream psychology in the twentieth century. The main reason for this is that the concept of image of oneself has not been grounded either in clearly determined psychological phenomena nor in morphological or functional brain structure (3). Nevertheless, the term 'image of the self' and others equivalent to it were broadly used in psychology of personality and social psychology, due to their practical convenience. The situation began to change in 1970s, when it became clear that the metaphors similar to the one in Figure 1 can be expressed in the language of functions and so provide formal description of human behavior. This opened the prospect of linking the human introspective world with objectively observable behavior (4, 5, 6, 7, 8, 9, 10, 11, 12).

This finding allows us to assume that the structure of the inner domain given in Figure 1 is a manifestation of work of a special cognitive mechanism of self-representation (possibly inborn), rather than a result of the intellectual efforts of the subject consciously thinking about the self (5, 6, 13).

We will demonstrate further how a model of the subject based on the reflexive metaphor allows us to suggest a unified explanation for three different psychological phenomena none of which has so far been explained convincingly.

A Function of Readiness

Let the subject face a choice between two alternatives: one of them plays the role of the positive pole for the subject and the other one that of the negative pole. Any bipolar opposition can serve as an example: good-evil, big-small, white-black (14, 15). Variable X_1 corresponds to the subject defined on the interval [0,1]. The value of this variable is called the subject's readiness to choose the positive pole. It can manifest itself in two ways: (a) as a frequency with which the subject chooses the positive alternative (under fixed condition); (b) as a dot on a scale [0,1] indicated by the subject and reflecting his readiness to choose the positive alternative. The subject is represented by the following function:

$$X_1 = x_1 + (1 - x_1)(1 - x_2)x_3, \qquad (1)$$

where $x_1,x_2,x_3 \in [0,1]$ (16). Variable x_3 describes the subject's intention to choose the positive pole. Thus, the model distinguishes between the objective readiness (X_1) of the subject's executive system to make a choice and his subjective desire to do so (x_3). Variables x_1 and x_2 describe the environmental pressure on the subject toward choosing the positive pole: x_1 is an actual pressure in the framework of a given situation, and x_2 is the subject's expectation of such pressure, as determined by his previous experience. Each alternative possesses a degree of attractiveness for the subject, expressed in the units of some psychological scale. The pressure toward the positive pole is connected with the alternative's attraction as $x_1=(v_1)/(v_1+v_2)$ and $x_2=(u_1)/(u_1+u_2)$, where v_1 and v_2 represent objective attraction of the positive and negative alternative in a given situation, and u_1 and u_2 show the expected attraction of these alternatives. In general, the alternative's attraction does not depend on its polarity. A negative alternative may be more attractive, as in a case of giving in to a temptation instead of rejecting it. Function Eq.1 can be represented as a composition $X_1=F(x_1,(F(x_2,x_3))$. This representation is unique and $F(x_2,x_3)=1-x_3+x_2x_3$ (16). Function $F(x_2,x_3)$ can be interpreted as the subject's 'image of the self'. Under this interpretation, variable x_3 is the image of the self in the image of the self. We will call this image of the second-order the 'model' of the self. We see that, in addition to the subject's intention, the value of x_3 plays the role of the readiness of the model of the self to choose the positive pole. The structure of composition $F(x_1,(F(x_2,x_3))$ corresponds to the metaphor in Fig.1.

The Law of Self-Reflexion

The manifest meaning of an intentional action is the subject's readiness to do what he intends to do. Variable X_1 corresponds to the subject's readiness, and variable x_3 to his desire. Thus, an intentional act corresponds to the condition $x_3=X_1$, where X_1 represents the subject as a whole, and x_3 represents his model of the self. The condition $x_3=X_1$ can be formulated as follows:

The subject tends to generate a pattern of behavior such that similarity is established and preserved between the subject and his model of the self.

For $x_3=X_1$, Eq.1 turns into

$$X_1 = \frac{x_1}{x_1 + x_2 - x_1 x_2}, \tag{2}$$

where $x_1+x_2>0$ (16). Let us note that the condition of similarity allows us to eliminate variable x_3, whose value is not instrumentally measurable.

Phenomenon 1. Non-linear Connection between Magnitude Estimation and Categorization of the Same Stimuli

Magnitude estimation is the choice of a number characterizing the intensity of a physical stimulus. For example, subjects are presented with a set of steel rods one by one and asked to estimate the length of each in inches. The data obtained from a large number of subjects allow experimenters to find a function G, which connects stimuli estimations with their objective physical measures. This function leads to the construction of a psychological scale of stimuli intensity. Categorization, on the other hand, classifies the stimuli according to their intensity. For example, the subjects are shown the same steel rods as before, but here the task is to refer each of them to one of eleven categories: the shortest rod belongs to the first category, the longest one to the eleventh category, and all the others lie in between (17). For a long time it was considered obvious that estimations obtained in these two kinds of experiments - magnitude and categorical - would be related linearly with each other. In the 1950's it was found that this relation is non-linear (17). It turned out also that the shape of a curve depends on the distribution of weak and strong stimuli in an experimental series: the more marked shift toward weak stimuli, the more convex the graph (18) (See Fig.2a).

Let us now connect these observations with the function of readiness (Eq.2). We represent a categorical scale as a segment [0,1], where the category of the strongest stimulus corresponds to point 1, playing the role of the positive pole, and the category of the weakest stimulus corresponds to point 0, playing the role of the negative pole.

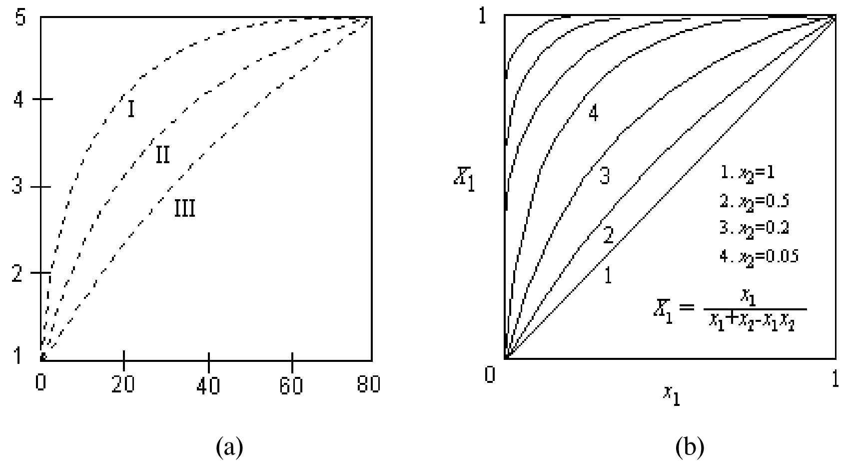


Fig. 2. Relation between magnitude estimation and categorization. (a) Categorization of rectangular areas (on five-point scale (16)): I - small stimuli more frequent; II - equal distribution; III - large stimuli more frequent. The vertical axis represents categorical scale; the horizontal axis corresponds to the area of rectangles measured in square inches (17). (b) A family of hyperbolas $X_1=(x_1)/(x_1+x_2-x_1x_2)$, where x_2 is a variable parameter (16).

All the other categories correspond to equidistant marks on this segment. We assume that the subject indicating point X_1 on the scale corresponds to a stimulus categorization (to within the nearest mark). Let G_{min} and G_{max} be the minimal and maximal intensities in a given experimental series (on the psychological scale). Then $v_1=G-G_{min}$ and $v_2=G_{max}-G$, where G is the intensity of the presented stimulus. It is easy to see that the greater G, the greater v_1 and the smaller v_2. Thus,

$$x_1 = \frac{v_1}{v_1+v_2} = \frac{G-G_{min}}{G_{max}-G_{min}}.$$

Let a given stimulus be preceded by a sequence of stimuli with the mean value of intensity G^*. Let also $u_1=G^*-G_{min}$ and $u_2=G_{max}-G^*$, then

$$x_2 = \frac{u_1}{u_1+u_2} = \frac{G^*-G_{min}}{G_{max}-G_{min}}.$$

For long randomized sequences of stimuli, x_2 will not change considerably after a series of presentations and can be considered constant. Under such conditions, Eq.2 turns into the equation of a hyperbola with variable x_1 and parameter x_2 (16). A family of such hyperbolas is shown in Fig.2b. Now we can explain the observations obtained during stimuli categorization:

(a) The connection between magnitude estimation and categorization is non-linear, because Eq.2 with constant x_2 corresponds to a hyperbola.

(b) The subjects overestimate the stimuli intensity in categorization by comparison with magnitude estimation, because hyperbolas are upward convex.

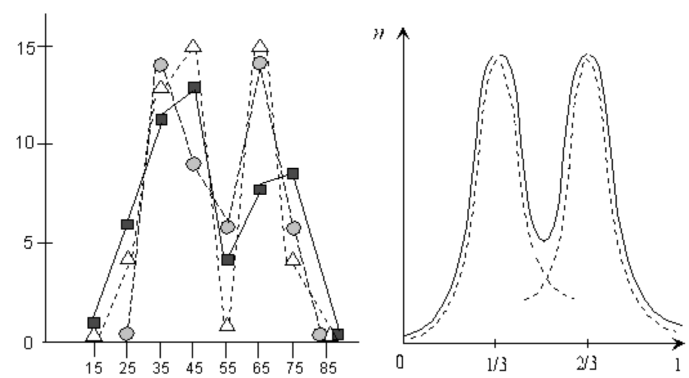
(c) When the intensity of stimuli is shifted toward the weakest values, the curve's convexity increases, because the value of parameter x_2 decreases.

The explanations currently existing for the entire set of the observations described above are based on the models containing a free parameter (19), while in the model based on Eq.2, there is no need for free parameters.

Phenomenon 2. Avoiding the Point 0.5 in Estimating Stimuli whose Intensity is Exactly in the Middle between Two Samples

This phenomenon was discovered by Poulton and Simmonds (20, 21, 22). The subjects were asked to evaluate the degree of lightness of a gray piece of paper situated close to two samples, one black and one white. The degree of lightness of the gray sample was chosen to be exactly in the middle between the black and white samples on a psychological scale. Each subject received a 100mm scale, one end of which corresponded to the black sample, and the other to the white one. The experimenter recorded only the very first touch of a pencil to the scale. A sample of experimental data is shown in Fig.3a: the graph has two peaks and a gap between them.

Let us link these experiments with Eq.2. Suppose that for one portion of the subjects, the white sample represents the positive pole, and for others, the black sample plays this role. The intensity of the gray sample lies exactly in the middle between the white and black ones, so that, $x_1=1/2$. Since only the very first touch of the pencil is registered, the subject's experience in making such estimations was limited to this one touch, so $x_2=x_1=1/2$. By substituting these values into Eq.2, we find that $X_1=2/3$. For the subjects, whose positive pole is the white sample, estimations will group at 2/3 of the distance from the left-hand end of the scale, i.e. around the point 2/3; for those, whose positive pole is the black sample, the estimations will group at the point located 2/3 of the distance from the right-hand end, i.e., around the point 1/3. As a result we obtain the two-humped distribution shown in Fig.3b.



(a) (b)

Fig. 3. Avoidance of the middle of the scale in estimation of stimuli intensity. The horizontal axis corresponds to a 100mm scale; the vertical one to the number of marks. (a) A sample of experimental distribution for three groups of subjects (20). (b) The distribution based on the reflexive model (16).

The only other explanation for this phenomenon assumes the existence of a shift of the scale's initial point, which is equivalent to introducing free parameters (21). There is no free parameters in Eq.2, with the help of which we explained this phenomenon, but there is the assumption that portion of the subjects has one orientation for the scale, and the other portion has an opposite orientation.

Phenomenon 3. Matching Law.

This phenomenon, discovered by Herrenstein in the early 1960s, manifests itself in experiments with birds and rats as well as in experiments with human beings (23, 24, 25). A standard experiment of this type proceeds as follows. Two keys connected with a food hopper are placed in a cage with a pigeon. A peck by the pigeon on either key may cause the appearance of a small piece of food. For each key, there is a special program of reinforcement which may be changed over the course of the experiment. For example, reinforcements may appear randomly, but a pigeon must make k pecks, on average, on one key to receive one grain. By varying the reinforcement programs for each key, experimenters succeeded in changing the frequencies of pecks by the pigeon on different keys. Analyzing the experimental data, Herrenstein found that the ratio $(N_2/N_1)=(n_2/n_1)$ holds; N_1 and N_2 being the numbers of pecks at each key, and n_1 and n_2 the numbers of reinforcements related to the corresponding keys. A more precise law was formulated by Baum in the middle 1970s:

$$\frac{N_2}{N_1} = p\frac{n_2^{\ S}}{n_1^{\ S}}, \tag{3}$$

where p and s are free parameters whose values may differ for different subjects (26, 27). In many experiments the value of s is close to 1.

Let us now link Eq. 3 with the function of readiness for the case of s=1.. Without compromising the generality of our reasoning, we choose the keys numeration in such a way that p#1. Eq.2 can be rewritten as

$$\frac{1-X_1}{X_1} = x_2\frac{1-x_1}{x_1}, \tag{4}$$

where x_1>0. Let $X_1 = \frac{N_1}{N_1+N_2}$, $x_1 = \frac{n_1}{n_1+n_2}$, and $x_2 = \frac{u_1}{u_1+u_2}$, where u_1 is an expected subjective attraction of the first key and u_2 that of the second. According to the reflexive model, in the first phase of the experiment there occurs a polarization of the keys (for each subject): one of them takes on the role of the positive pole, the other that of the negative pole. At the same time, the expected attractions, u_1, u_2, are formed, as well as the value of x_2 which plays the role of p in Eq.3. We see that the law of reflexivity allows us to obtain Eq.3 and provide parameter p with a psychological interpretation for the case when s=1 (28).

Unlike phenomena 1 and 2, the matching law is not connected with the subject's estimating activity, but rather with his economical behavior. It seems natural to assume that this law reflects the tendency of an organism to obtain as much utility as possible. This idea underlies most attempts to explain this phenomenon, although Heyman and Luce demonstrate that the matching law is not a logical consequence of maximizing reinforcement rate (29). Nevertheless, many researcher do not rule out the possibility that the subject seeks to maximize utility understood in broader sense (30). For example, the subject may try to shorten the run between a key and the food hopper or save the energy needed to operate with the key, ans so forth. Baum and Apparacio mention this topic, "Despite claims to the contrary, all leading theories about operant choice may be seen as model of optimality." (31, p.75). The idea of maximizing utility values, however, has not helped researchers to deduce Eq.3 (32).

The reflexive model provides us with another possible explanation for the matching law: Eq.3 holds not because the subject tends to obtain more utility of any kind, but because he generates a pattern of behavior such that the relation of similarity between the subject and his model of the self is established and sustained.

If the reflexive model could explain *why* it is necessary to introduce a free parameter *s* into Eq.3, this would become an important step forward toward substantiating this hypothesis.

References and Notes

1. J. Locke. *An Essay Concerning Human Understanding* (1690). Amherst: Prometeus Books (1995).
2. E. Cassirer. *The Philosophy of the Enlightenment.* Princeton NJ: Princeton University Press (1951).
3. P. S. Churchland. Self-Representation in Nervous Systems. *Science*, **296**, 308 (2002).
4. V. A. Lefebvre. A Formal Approach to the Problems of Good and Evil. *General Systems*, **22**, 183 (1977).
5. V. A. Lefebvre. An Algebraic Model of Ethical Cognition. *Journal of Mathematical Psychology*, **22**, 83 (1980).
6. V. A. Lefebvre. *Algebra of Conscience*, Dordrecht: D.Reidel (1982). Second enlarged edition, Dordrecht: Kluwer (2001).
7. J. Adams-Webber. Comment on Lefebvre's Model from the Perspective of Personal Construct Theory. *Journal of Social and Biological Structures*, **10**, 177 (1987).
8. J. A. Schreider. Fuzzy Sets and the Structure of Human Reflexion. *Applied Ergonomics*, **1**, 19 (1994).
9. V. Yu. Krylov. On One Model of Reflexive Behavior Distinct from Lefebvre's Model. *Applied Ergonomics*, **1**, 21 (1994).
10. J. Adams-Webber. Self-Reflexion in Evaluating Others. *American Journal of Psychology*, **110**, 527 (1997).
11. L. D. Miller & M. F. Sulkoski. Reflexive Model of Human Behavior. *Proceedings of Workshop on Multi-Reflexive Models of Agent Behavior*, Los Alamos: Army Research Laboratory (1998).
12. T. A. Taran. Many-Valued Boolean Model of the Reflexive Agent. *Multi-Valued Logic*, **7**, 97 (2001).
13. V. A. Lefebvre. The Golden Section and an Algebraic Model of Ethical Cognition. *Journal of Mathematical Psychology*, **29**, 289 (1985).
14. G. A. Kelly. *The Psychology of Personal Constructs*, New York: Norton (1955).
15. C. E. Osgood, G. J. Suci, & P. H. Tannenbaum. *The Measurement of Meaning*, Urbana, IL: University of Illinois Press (1957).
16. V. A. Lefebvre. *A Psychological Theory of Bipolarity and Reflexivity*, Lewiston: The Edwin Mellen Press (1992).
17. S. S. Stevens & E. H. Galanter. Ratio Scales and Category Scales for a Dozen Perceptual Continua. *Journal of Experimental Psychology*, **54**, 377 (1957).
18. A. Parducci. Direction of Shift in the Judgment of Single Stimuli. *Journal of Experimental Psychology*, **51**, 169 (1956).

19. A. Parducci. Category Judgment: A Range-Frequency Model. *Psychological Review*, **72**, 407 (1965).
20. E. C. Poulton, D. C. V. Simmonds, R. M. Warren. Response Bias in Very First Judgments of the Reflectance of Grays. *Perception & Psychophysics*, **3**(2A), 112 (1968).
21. E. C. Poulton & D. C. V. Simmonds. Subjective Zeros, Subjectively Equal Stimulus Spacing, and Contraction Biases in Very First Judgments of Lightness. *Perception & Psychophysics*, **37**, 420 (1985).
22. E. C. Poulton. *Bias in Quantifying Judgments*, NJ: Erlbaum (1989).
23. R. J. Herrnstein. Relative and Absolute Strength of Response as a Function of Frequency Reinforcement. *Journal of the Experimental Analysis of Behavior*, **4**, 267 (1961).
24. G. M. Heyman & R. J. Herrnstein. More on Concurrent Interval-Ratio Schedule: A Replication and Review. *Journal of the Experimental Analysis of Behavior*, **46**, 331 (1986).
25. B. A. Williams. Reinforcement, Choice, and Response Strength. In: R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Steven's Handbook of Experimental Psychology* (Vol.2), New York: John Wiley & Sons (1988).
26. W. M. Baum. On Two Types of Deviation from the Matching Law: Bias and Undermatching. *Journal of the Experimental Analysis of Behavior*, **22**, 231 (1974).
27. W. M. Baum. Matching, Undermatching, and Overmatching in Studies of Choice. *Journal of the Experimental Analysis of Behavior*, **32**, 269 (1979).
28. V. A. Lefebvre. Categorization, Operant Matching, and Moral Choice. Institute for Mathematical and Behavioral Sciences, *MBS*, **99-14**, UCI (1999).
29. G. M. Heyman & R. D. Luce. Operant Matching Is Not a Logical Consequence of Maximizing Reinforcement Rate. *Animal Learning Behavior*, **7**, 133 (1979).
30. D. W. Stephens & J. R. Krebs. *Foraging Theory*, Princeton, NJ: Princeton University Press (1986).
31 W. M. Baum & C. F. Aparicio. Optimality and Concurrent Variable-Interval Variable-Ratio Schedule. *Journal of the Experimental Analysis of Behavior*, **71**, 75 (1999).
32. G. M. Heyman. Optimization Theory: A Too Narrow Path. *Behavioral and Brain Sciences*, **11**, 136 (1988).