The Economics of Curiosity

Jeong-Yoo Kim*
Kyung Hee University

Haeree Lee
Kyung Hee University

Insik Min
Kyung Hee University

July 21, 2009

Abstract

We develop the hypothesis that an individual can get some value of information, even if they do not use the information for his subsequent decision, contrary to the expected utility theory. Curiosity is associated with the direct utility from information and is defined formally by using the concept of entropy. We can measure an agent's curiosity level by the maximum amount of money that he is willing to pay in order to obtain the information thereby reducing the entropy. We test the hypothesis from lab experiments and obtain the empirical evidence that people are actually willing to pay a positive amount of money to obtain payoff-irrelevant information. Also, the comparison of the coefficients of variation for our curiosity measure and the IPI curiosity measure which is widely used in psychology suggests that our measure is more informative.

Key Words: Curiosity, Entropy, Expected Utility, Shannon Value

PsycINFO Classification: 2224

JEL Classification: K42

*Corresponding author: Department of Economics, Kyung Hee University, 1 Hoegidong, Dongdaemunku, Seoul 130-701, Korea, (Tel) +822-961-0986, (Fax) +822-961-0622, (Email) jyookim@khu.ac.kr

1 Introduction

Many people obtain utility from learning something they did not know. Some often pay significant costs simply to learn something they want to know. For example, we have fun for quiz games, are excited by news about a movie star, have interest in the culture of other countries, enjoy reading about history, science, art etc. or observing flowers, insects, stars, etc. We sometimes even struggle to find out a solution for a puzzling mathematical problem without any reward, and some explore the unknown land at the risk of the life. These behavior can be explained by the notion of "curiosity."

In this paper, we define curiosity in a formal way. In plain English, curiosity is defined by someone's propensity to learn something unknown to him.[1] The motive of learning something may be various, but we distinguish mainly between two motives; one is to avoid risks associated with an uncertain monetary outcome and the other is purely to want to know it without any monetary concerns.[2] In the former case, information gives indirect utility, while it gives direct utility in the latter case. We will call only the latter motive curiosity. For instance, if a person is eager to know the answer for a quiz question "What is the oldest existing firm in the world?" or "What is the highest mountain in Europe?," it must be out of curiosity, but if a person who is thinking of investing in stocks wants to know "What will be the interest rate next month?," it is not.

Curiosity is a personal characteristic, so the intensity of curiosity can vary across individuals.[3]

[1]Litman (2005) defines curiosity by the "desire to know, to see or to experience leading to exploratory behavior directed towards the acquisition of new information." Here, seeing and experiencing are both to acquire new information, i.e, to know it. Therefore, in a broad sense of the word "know", the phrase "to see or to experience" may be considered as redundant, although psychologists distinguish sensory curiosity (perceptual curiosity) for sensing experience from cognitive curiosity (epistemic curiosity) for knowing experience. See Loewenstein (1994) and Litman and Spielberger (2003).

[2]Condry (1977) distinguishes curiosity as "intrinsic" and "extrinsic." While extrinsic curiosity is to receive an external reinforcement, intrinsic curiosity is independent of external reward, such as play, imaginative behavior etc. Thus, intrinsic curiosity corresponds with our latter motive. Also, Wohlwill (1981) calls it *affective* curiosity (or exploration) whose examples include children's high level play for the pure joy of it and adults' philosophizing.

[3]It is reported that the intensity of curiosity is both inherited and acquired. While Saxe and Stollak (1971) found some evidences for the social learning theory that both parental reinforcement and modeling foster children's curiosity, the characteristic of curiosity is recently found in genomes. See Fidler *et al.* (2007).

Insofar as curiosity is associated with an individual's utility obtained from some information, it is natural that each individual attaches a different value to given information. Then, what can be used as an index for an individual's curiosity? How can we measure his valuation for information? For the purpose, the information theory can provide a useful index of curiosity.

Entropy is a measure of disorder or a measure of the uncertainty. It quantifies information (usually in bits). For example, a fair coin has an entropy of one bit. However, if an agent knows that the outcome of the coin toss is Head, the entropy is zero, because there is no further information necessary to communicate. A fair dice has a higher entropy than a fair coin, because it has more equally likely outcomes.

If an agent does not know the value of a random outcome, curiosity indicates that he wants to know it. This means that he has a valuation for the information. Note that the value of information can be measured by entropy.[4] This implies that learning the value of a random outcome by obtaining information is a process of reducing entropy,[5] and naturally we can conclude that *curiosity is a biological mechanism to lower entropy.*[6]

Learning an unknown value can be done by spending some resources in acquiring the information. The more curious he is of the random outcome, the more resources he will be willing to invest. Thus, we can use the maximum amount of money that an agent is willing to pay in order to obtain information as the proxy for his intensity of curiosity. In analogy with risk preference, we will say that an individual is curious if he is willing to pay a positive amount of money to obtain given information, or otherwise he is incurious.[7] Also, we will say that an individual is more curious than another if he is willing to pay more to acquire given information.

---

[4]The attempt to interpret entropy as the value of information was begun by Kelly (1956) and then has been succeeded to Bellman and Kalaba (1957), Marschak (1959) and Arrow (1972).

[5]The view of curiosity from an information-gap perspective dates back to William James (1890 [1950]). Also, see Kreitler, Zigler, and Kreitler (1974) for an interesting view on the entropy-interpretation of curiosity.

[6]Erwin Schrödinger (1944), in his famous book "What is Life?", used the concept of negative entropy. By the concept, he meant that a living system exports entropy in order to maintain its own entropy as a low level. It is an interesting analogy that information entropy is reduced by a flow of energy which can be initiated by curiosity.

[7]This case includes that an individual is reluctant to know something, i.e., he even wants to pay a positive amount of money in returns for leaving him ignorant of the information. For example, some patients may not want any information about their fatal medical conditions.

If it is too costly to obtain full information, people may alternatively want to acquire noisy, partial information in a less costly way in order to satisfy their curiosity at least partially thereby reducing their entropy. The information theory predicts that people will be willing to pay more to obtain a more informative signal for the information.

To test the hypothesis that people may get positive utility simply from learning the value of a payoff-irrelevant random outcome, we perform a laboratory experiment. After recruiting about 40 subjects in a random and public way, we provide 40 questions which may arouse their curiosity and ask them to respond by mouse-clicking which is costly in terms of their reward they will be paid after the experiment. The experimental results suggest that people are actually willing to pay a positive amount of money to obtain payoff-irrelevant information, and that their willingness-to-pay (WTP) for full information is on average higher than their WTP for partial information. Both results seem to support the hypothesis and the explanation of curiosity based on entropy. We also find that the curiosity level of an individual for persons (interpersonal questions) is positively correlated with his curiosity level for objects (impersonal questions). In addition, we compare our curiosity measure (WTP) of an individual with Imaginal Processes Inventory (IPI) measure which has been widely used in psychology[8] since Singer and Antrobus (1970) invented it. The distributions of WTP and IPI over subjects seem to suggest that our measure is more informative than IPI measure in the sense that the former tends to have a higher coefficient of variation than the latter, that is, its distribution is more disperse across individuals. Interestingly, our experimental result also show that the difference in curiosity level between males and females is rejected in all cases at a 5% significance level.

While curiosity is considered as one of the most important motivations for learning in schools, little about its effect on the economic behavior has been explored so far. However, curiosity indeed has a deep potential for explaining lots of human (economic/social) behavior that has been unexplored. For example, many youngsters start adolescent deviance including

[8]Psychologists use measures of curiosity which are outcomes of responses to pencil-and-paper questionnaires. They include the Novelty Experiencing Scale (NES; Pearson, 1970), the Academic Curiosity Scale (ACS; Vidler and Rawan, 1974), the Sensation Seeking Scale (SSS; Zuckerman, 1979, 1994) and the Melbourne Curiosity Inventory (MCI; Naylor, 1981) etc. For example, Zuckerman (1994) contains items like "I would like to try parachute jumping."

smoking and drug use simply out of curiosity.[9] In South Korea, it is reported that "almost one in every two female smokers was found to have started smoking out of simple curiosity."[10] This implies that their choices are made heavily based on the consideration of curiosity rather than simply based on expected utility. Also, gossip magazines featuring scandalous stories about the personal lives of celebrities have been flourishing. Most of such magazines are designed mainly for catering to the curiosity of people. These days, there are nearly 400 magazines including *People* and *Us* related to gossip sold in the U.S. news stand.

The paper is organized as follows. In Section 2, we formally introduce the concept of information entropy and define curiosity in terms of entropy. In Section 3, we consider the effect of noisy information to satisfy curiosity partially. In Section 4, we test the hypothesis that people can direct utility from information by using laboratory experiments. In Section 5, we discuss some economic applications. In Section 6, we compare our theory with the expected utility theory. In Section 7, we briefly discuss the extended decision-making models. Concluding remarks follow in Section 8.

2 Entropy and Curiosity

Let X be a random variable. If we want to define uncertainty $H(X)$ associated with X, $H(X)$ is to be a function only of the probability distribution of X. In 1948, Shannon proposed a concept of information entropy as a measure of uncertainty. He started from three axioms that a sensible definition of uncertainty $H(X)$ should satisfy;

Let X be a random variable with values x_i and probabilities $p_i = \text{Prob}(X = x_i)$, $i = 1, \ldots, n$. Then, $H(X)$ must satisfy

Axiom 1 H is continuous in p_i.

Axiom 2 If all the p_i are equal, $p_i = 1/n$, then H is a monotonic increasing function of n.

Axiom 3 H of a composite choice is the weighted sum of the individual values of H.

[9] See Green (1985), Mizner *et al.* (1970) and Ormian (1975).

[10] See Byun (2003).

Then, he established the following well known theorem.

Theorem 1 (Shannon) *The only H satisfying Axiom 1 – 3 is of the form*

$$H(X) = -k \sum_{i=1}^{n} p_i \log p_i, \tag{1}$$

where k is a positive constant.

Proof. See Shannon (1948).

$H(X)$ is called Shannon value or information entropy.

Corollary 1 *$H(X) \leq -k \log n$ with equality if and only if $p_i = 1/n$ for all i.*

Proof. See the appendix.

This corollary gives the maximum value of entropy. Suppose one tosses a fair coin and tell you the outcome *Head* of the experiment. Now, suppose one tosses a biased coin and tell you the outcome *Head* when the probability of *Head* is 0.99. Then, in the latter case, considerably less information is provided than in the former case, since the outcome is already expected. This example suggests that the information of an event $X = x_i$ should be defined as

$$I(X = x_i) = -\log p_i. \tag{2}$$

This implies that the entropy of a random variable is the expected value of the information content of events $X = x_i$. Therefore, we can say that entropy measures the average information content of an observation of X.

Now, we can define curiosity in terms of entropy. Consider an agent and a variable X that yields an uncertain outcome but whose value is irrelevant to his monetary income. We normalize the utility when he knows the value of a random outcome to zero, i.e., $\bar{U} = U(X = x_i) = 0$ for any x_i. If he does not know the value, this yields him disutilty. Our hypothesis is that this disutility can be measured by the entropy, that is, $U = U(X) = -H(X) = k \sum_{i=1}^{n} p_i \log p_i$. Assume that the utility is the (monetary) value to the agent, accordingly measured in the monetary unit. Then, his value of the information $(X = x_i)$ is $\Delta U \equiv \bar{U} - U(X) = H(X)$. Therefore, this agent will be willing to pay the price up to

$p = \Delta U = H(X)$ to obtain the information. This implies that $p = H(X)$ can be viewed as a measure of curiosity.[11] The maximum amount of money that an agent is willing to pay to eliminate uncertainty (not risk) can be called *uncertainty premium*. It has the exact analogy with risk premium in the case of payoff-relevant information. Also, given X, $\Delta U = H(X)$ can differ across agents. We can think of the difference as coming from a difference in k. In fact, high $H(X)$ is associated with high k. Thus, k can be also used as a measure of curiosity.

To elaborate, consider two agents A and B. Let U_A and U_B be the utility levels of the respective agents when they have no information. If $U_A < U_B$, it means that $k_A > k_B$, or equivalently, $H_A(X) > H_B(X)$, in turn implying that agent A is willing to pay a higher price than agent B to learn the value. Thus, we can say that agent A is more curious than agent B. If $\Delta U = H(X) = 0$ for some agent, that is, he wants to pay nothing to learn the value of X, we can say that he is incurious, which corresponds to the case that $k = 0$.

3 Noisy Signal

When an agent wants to know the value of X, he may have a chance to obtain a noisy signal of X instead of directly learning the value of X.

Let Y be a noisy signal of X with values y_j for $j = 1, \ldots, m$ and $P(X = x_i \mid Y = y_j)$ be the conditional probability. Unless X and Y are independent, i.e., $P(X = x_i \mid Y = y_j) = P(X = x_i)$, we can say that Y is informative. Then, the conditional entropy of X given $Y = y_j$ can be defined by

$$H(X \mid Y = y_j) = -\sum_{i=1}^{n} P(X = x_i \mid Y = y_j) \log P(X = x_i \mid Y = y_j)$$

and the conditional entropy of X given Y is defined as the weighted average of the entropies $H(X \mid Y = y_j)$ for $j = 1, \ldots, m$, i.e.,

$$\begin{aligned} H(X \mid Y) &= \sum_{j=1}^{m} P(Y = y_j) H(X \mid Y = y_j) \\ &= -\sum_{j=1}^{m}\sum_{i=1}^{n} P(Y = y_j) P(X = x_i \mid Y = y_j) \log P(X = x_i \mid Y = y_j). \end{aligned}$$

[11] We are implicitly assuming that all agents get the same utility from money.

Conditional entropy measures the average uncertainty of a random variable X given observations of a random variable Y, averaged over all values that Y can take.

Let us take a simple example of $n = m = 2$. Suppose X can take only two values of x_1 and x_2 with equal probabilities. Let Y be a noisy signal of X, so that if $Y = y_1$, $X = x_1$ with probability $\rho(\geq 1/2)$ and if $Y = y_2$, $X = x_1$ with probability $1 - \rho$. If $\rho = 1/2$, X and Y are independent, and as ρ increases, Y is a more informative signal of X. We know that $H(X) = k \log 2$. Now, compute $H(X \mid Y)$. We have

$$H(X \mid y_1) = H(X \mid y_2) = -k \left[\rho \log \rho + (1-\rho)\log(1-\rho)\right],$$

implying that

$$H(X \mid Y) = -k \left[\rho \log \rho + (1-\rho)\log(1-\rho)\right].$$

Let $\psi(\rho) \equiv \rho \log \rho + (1-\rho)\log(1-\rho)$. Note that $\psi'(\rho) = \log \frac{\rho}{1-\rho} > 0$. This means that $H(X \mid Y)$ is decreasing in ρ. In other words, $H(X \mid Y)$ becomes lower as Y is a more informative signal of X. To generalize, we have

Theorem 2 *$H(X \mid Y) \leq H(X)$ with equality if and only if X and Y are independent.*

Proof. See the appendix.

Note that this is true averaging across all possible values of Y. It does not necessarily hold for every realization of Y, i.e., it is not always the case that $H(X \mid Y = y_j) \leq H(X)$ for all j.

How much is an agent willing to pay to acquire the value of a signal Y? The gain he obtains by paying the price is $\Delta = H(X) - H(X \mid Y)$, which must be the maximum price that he is willing to pay. If X and Y are independent, Y is not informative at all, so he will not pay any positive price. As Y is more informative, he will be willing to pay more. Also, as far as $H(X)$ and $H(X \mid Y)$ are linear in k, a more curious agent will be willing to pay more to have access to a given signal.

Now, let us formally define more informativeness. Let Y and Z be two signals of X where $|Y| = m$ and $|Z| = l$. Also, let $P_Y = [P(Y = y_j \mid X = x_i)]$ $n \times m$ (conditional probability) matrix and $P_Z = [P(Y = z_h \mid X = x_i)]$ $n \times l$ matrix each of which specifies a probability distribution over a set of signal values for each x_i. Then, à la Blackwell (1951), we can say that Y is more informative than Z if there exists an $m \times l$ matrix M with $P_Y M = P_Z$. The matrix M plays the role of stochastically transforming Y to Z, and it must itself satisfy the usual conditions of a conditional probability distribution in the sense that the entries in each row sum up to one. In this case, Z is referred to as a "garbling"[12] of Y since it is as if Z were generated from Y using a stochastic transformation. Each realization z_h can be interpreted as being obtained from Y by adding some noise through a process of randomization. Now, we have

Theorem 3 *If Y is a more informative signal of X than Z, then $H(X \mid Y) < H(X \mid Z)$.*

Proof. See the appendix.

This is intuitively clear. If one can obtain a better information generator in the sense that it yields a more informative signal, he can reduce the entropy more. So, he will be willing to pay a higher price for the signal to learn the value of X.

4 Experimental Evidences

In this section, we test the hypothesis that people obtain utility from the probability distribution as well as from the reward by using laboratory experiments. Since they will be willing to pay nothing if and only if they obtain no such utility, it is equivalent to show that they are willing to pay some positive price to obtain information even if the information does not affect their monetary income. Also, we use an individual's willingness-to-pay (WTP) as a measure of curiosity, and compare it with Imaginal Processes Inventory (IPI) curiosity measure which is widely used in psychology. IPI, which was first developed by Singer and Antrobus (1970), consists of 28 subscales and 4 of them are to measure curiosity motivations. Among them, two subscales (Interpersonal Curiosity and Impersonal Curiosity) are to measure information-seeking motives of curiosity, while the other two subscales (Boredom and Need for External Stimulation) are to measure the stimulation-seeking motives. For our purpose, we used only the two subscales measuring information-seeking motives and selected

[12]Marschak and Miyasawa (1968) used this term of "garbling."

10 questions each from interpesonal questions and impersonal/mechanical questions.[13] The Cronbach α, which is a measure of internal consistency for a set of questions, was 0.559 for interpersonal questions and 0.774 for impersonal questions.[14]

4.1 Experimental Design

The experiments were conducted at the Laboratory of the economics department at Kyung Hee University on October 8, 2008. Subjects were recruited by the announcements in the community website in Kyung Hee University and Hankuk University of Foreign Study which is located within a walking distance from Kyung Hee University. The announcements were posted on October 3, 2008. We pre-excluded the students from the economics department of Kyung Hee University to avoid a possible demonstration effect.[15]

The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). 42 students participated in the experiment. To make participants clearly understand the experimental rules, we explained the instructions one more time by illustration after asking them to read the instructions carefully.

To test that subjects are willing to pay some monetary amount to obtain full or partial information regarding uncertain outcomes, we used a set of 40 questions.[16] In Part A consisting of 20 questions, subjects are supposed to choose whether or not they are interested in obtaining the full answer for each question at the price of 100 won (roughly 10 cents). We set this price by considering the price for the alternative information source which is the information usage fee in the Internet. If they choose to obtain the full answer, the correct answer is provided; otherwise, it is not provided. In part B consisting of the remaining 20 questions, they are supposed to choose whether or not they are interested in obtaining

[13]Some of impersonal questions were slightly modified or newly invented by the authors (#3, 9, 13, 15, 20), since many of the questions in this scale are more or less mechanical ones that have little to do with curiosity for objects, for example, "I have always liked to take things apart to see what makes them work" or "I know relatively little about the mechanical operation of an automobile."

[14]As a rule of thumb, social scientists say that the consistency of the questions is reliable if the value of Cronbach α exceeds 0.7 and that it is acceptable if the value is about 0.6 or higher.

[15]By the demonstration effect, we mean the psychology that students want to look good to their professor.

[16]One author made 142 questions, and the other two authors selected 40 questions of which they were most curious. This selection procedure was to maximize the entropy value for the regression purpose.

hints (noisy information) about the correct answer at the price of 50 won. If they choose to see the hints, two possible candidates for correct answers are provided.[17] In this part, the correct answer is never provided. We allocated 10 interpersonal questions and 10 impersonal questions in each Part in a random order so that the subjects cannot realize that they are being tested on curiosity for interpersonal matters or impersonal matters.

When a subject chose not to obtain information, it could mean either "he/she is really uninterested" or "he/she already knows the answer." To distinguish between the two and to measure the pure curiosity intensity, we asked subjects to provide the answer in the beginning of each question. Only when he/she failed to provide the correct answer, was it programmed so as for him/her to proceed the binary choice.

On average, the experimental session lasted about 40 minutes. The participants were paid their respective earnings including the showup fee of 15,000 won ($15) and the initial endowment of 3,000 won ($3) on October 9 after we completed the calculations of their earnings.

4.2 Empirical Results

We examine the correlation among several measures of curiosity for interpersonal questions (persons) and impersonal questions (objects) respectively from individual data. Then, by averaging the values of individual measures across subjects, we show that the willingness-to-pay (WTP) for payoff-irrelevant information, that is, people obtain direct utility from information which does not yield any monetary reward. Also, we examine the correlation between the given curiosity measures for interpersonal questions and impersonal questions, and compare those measures of males and of females.

Table 1 provides the summary statistics. To see the correlation between WTP for full information and WTP for partial information, we plot the pairs of the two WTP as in Figure 1. Similarly, to see the correlation between WTP for full information and IPI measure, we plot the pairs as in Figure 2. As Figure 1 suggests, there seems to be a strong correlation between WTP for full information and WTP for partial information both in cases of inter-

[17] The binary choices were provided as a pair of one correct answer and the other which most students are likely to choose. This is again to maximize the entropy value.

personal questions and impersonal questions. This correlation is summarized in Table 2. On the other hand, Figure 2 suggests that the correlation between WTP for full information and IPI measure seems to be weak or nonexistent. The correlation is also summarized in Table 2. Figure 3 shows the correlation between an individual's curiosity for interpersonal questions and curiosity for impersonal questions, and it is summarized in Table 3.

Now, let us denote the average curiosity level over all subjects by μ_{WTP}. Then, for our purpose, we consider the following null hypothesis; $H_0 : \mu_{WTP} = 0$. A hypothesis test shows that the null hypothesis is rejected at a 5% significance level, that is, people are willing to pay a positive amount of money to obtain payoff-irrelevant information. This result is summarized in Table 4. Then, to compare the FTP for full information and that for partial information, we denote the respective WTP by μ_{Full} and $\mu_{Partial}$. Since we obtained two observations for given subject, we use paired sample t-test. Table 5 shows that people are willing to pay a higher price for full information than for partial information. This result is consistent with Theorem 2.

Also, out of simple curiosity, we test various hypotheses that males and females are equally curious. Although Table 6 seems to suggest that males are more curious for both interpersonal questions and impersonal questions, interestingly none of the hypotheses is rejected at a 5% significance level.[18]

Figure 4 shows the distributions of various curiosity measures. Figure 4-a suggests that some people have extremely high curiosity for interpersonal questions, while all people have moderate curiosity for impersonal questions. Also, Figure 4-a and 4-b show that the distributions are all long-tailed, right-skewed which is different from the normal distribution. On the contrary, IPI measures seem to follow a normal distribution as shown in Figure 4-c.

Finally, we employ the coefficient of variation (CV) for the dispersion comparison. This statistics is to compare dispersion of two distributions when the means differ substantially. In our experiment, the means of WTP and IPI measures are so different that it is not appropriate to compare the dispersions of the two measures simply by standard deviations. As Table 7 shows, the dispersion of IPI is relatively small. Most of the subjects chose values around 3. On the other hand, the CV of WTP has the value which is more than 10 times

[18]Only one hypothesis that their WTP for partial information are same is rejected at a 10% significance level.

larger than that of IPI. That is, the curiosity index measured by WTP is much more disperse. This implies that WTP can be considered as a more informative measure of curiosity than IPI.

5 Economic Applications

In this section, we apply this line of thought to human economic behavior and attempt to understand the behavior that could not be explained within the framework of the traditional expected utility theory.

Curious-Related Market

People keep watching soap operas because they are just curious of what will be the next story. Children could not stop reading Harry Potter series simply because of curiosity.

Besides, clearly there is a market for information that can satisfy people's curiosity, for example, newspaper, gossip magazines, books, films etc. People are willing to pay prices in returns for obtaining the information they wanted out of curiosity. As the information provided by this channel is more accurate, people will be willing to pay more for the information channel. That is, newspaper, gossip magazines and so on can serve as a noisy signal. However, the market often goes further. There are some businesses whose tactic is to pique curiosity. In our framework, this tactic can be interpreted as giving people the misleading impression that there are more plausible possibilities than they thought. This can increase the entropy so as to induce them to pay more for their information products.

Curiosity-Inducing Marketing

In Korea, the placard saying "I love you, Sunyoung." that appeared from street to street in 2000 made people curious about what it is for, and after all, it turned out to be an advertising for www.miclub.com. This is an example of so-called teaser marketing. Teaser advertising sends cryptic ad messages by veiling the brand or deliberately omitting the explanation about the brand in order to induce consumers' curiosity. A typical example is Sony's ad campaign for Rolly, a commercial featuring a man dancing with the Rolly device, which was initially launched on August 20, 2007. Also, the number marketing that uses numbers in the name of the brand so as to arouse consumers' curiosity for what the numbers mean or the initial marketing that expresses the brand only by its alphabetical initials are often in use. The common feature of such marketing techniques is to pique curiosity about the advertising brand so as to give consumers a stronger desire to try the brand.

Only in the thin marketing literature has the effect of the marketing technique arousing curiosity been discussed. Krugman (1965) asserted that the curiosity-inducing advertising has the effect of making consumers search for information, possibly leading to purchase. Menon and Soman (2002) argued that consumers tend to remember the curiosity-inducing brand better.

Policy to Reduce Curiosity

The juvenile delinquency like smoking and drug use is often started out of simple curiosity. This is partly because a certain class of behavior is too strictly regulated, which makes the value of information too high. As a consequence, juveniles tend to experience such behavior even at the expense of significant risks or monetary prices.

If the government believes that it is socially desirable to discourage such behavior caused mainly by curiosity, the policy can be either to raise the cost of engaging in the behavior even higher, for example, by increasing the penalty, or to reduce the value of information. The latter policy can be enforced by offering chances to experience other alternatives yielding noisy signals. Since this is similar to boosting immunity to a specific disease by introducing a vaccine or a serum into the human body, it is called the *inoculation effect* in social psychology. For example, the sex education can have the inoculation effect of reducing the adolescence's high curiosity about sexes. The inoculation technique, which was first discovered by Lumsdaine and Janis (1953) and formulated by Papageorgis and McGuire (1961), can be also effective in preventing adolescent smoking. McAlister *et al.* (1980) reports that middle schools which offer the class of teaching the good sides and bad sides of smoking has much lower smoking rates than those schools which do not.

6 Relation with the Expected Utility Theory

Measuring the value of information has been studied by many authors including Kelly (1956), Bellman and Kalaba (1957), Marschak (1959), La Valle (1968), Hirshleifer (1971), Arrow (1972) and Murota (1988) etc.

To the best of our knowledge, all of them measure the value of information in terms of the difference between the maximum expected payoffs resulting from the decision maker's optimal decision when he is informed of the information and not. So, their models all implicitly assume that the decision maker attaches some value to the information to the extent that it increases its expected payoff, implying that their measure of the value of information is based on the expected utility theory. In our model, however, an individual gets the value of the information not because it helps him make a better decision thereby increasing his expected utility, but because he enjoys acquiring the information itself.

We will take Arrow's model to elucidate the contrast. Arrow considers an individual who bets on the occurrence of states of nature. As before, let p_i be the prior probability that $X = x_i$ for $i = 1, \dots, n$, and a_i be the amount bet on the occurrence of state i. Also, let $u(x)$ be his utility function where x is his monetary income. Assume that he invests all his resources which is equal to 1. The individual will then face the problem:

$$\max_{a_i} \sum_i p_i u(a_i x_i)$$
$$\text{s.t.} \sum_i a_i = 1.$$

If we assume $u(x) = \log x$, the optimal bet is $a_i^* = p_i$ and thus the maximized utility is $U_0 \equiv \sum_i p_i \log p_i + \sum_i p_i \log x_i$. On the other hand, if he knows the state of world, for example, by installing (purchasing) an information generator which shows a perfect signal, he can bet all his resources on the true value. This will yield him the utility of $\log x_i$, thus his expected utility is $U_1 \equiv \sum_i p_i \log x_i$. Therefore, the value of the information (signal) is

$$\Delta = U_1 - U_0 = -\sum_i p_i \log p_i = H(X).$$

Note that Arrow's interpretation of $H(X)$ as the value of information is possible only under the assumption that the utility function is logarithmic, while our interpretation requires no such assumption.

Since we empirically confirmed that people get direct utility from information which is not captued in the expected utility theory, we argue that the proper form of the utility function under uncertainty should be modified, and furthermore, we propose the following generalized expected utility function;

$$U(X) = \sum_{i=1}^{n} p_i u(x_i) + k \sum_{i=1}^{n} p_i \text{log} p_i. \tag{3}$$

The first term is the indirect utility from information and the second term is the direct utility from information (disutility from lack of information). Note that the proposed utility function *cannot be unique up to affine transformations.*

7 Discussion on the Extended Models of Decision-Making

In defining curiosity, we only compared the utility before and after uncertainty is resolved and paid little attention to the mechanism through which curiosity is resolved. In fact, there are various ways and degrees in which uncertainty is resolved. To see the picture, it is worthwhile to recognize that resolving uncertainty entails some cost of efforts. The effort can be to try either to resolve a problem for himself or to find someone who can help him to resolve it. Also, uncertainty may be resolved either completely or only partially as in our experiments. Obviously, the outcome will depend on the complication level and the cost of effort required to resolve the problem (uncertainty) completely.

Decision to Cope with Uncertainty When a new uncertainty arises, not everyone tries to resolve it. He may either approach or avoid/ignore it, depending on the complication level and the resolution cost.

The complication level refers to how uncertain the random outcome is, that is, how much more information is needed to resolve the uncertainty. Thus, the complication level and the entropy are conceptually identical, but if we use different units for them, we can assume that the entropy level is proportional to the complication level and that the efforts required to resolve also increases in the complication level. Let the complication level be θ and a random variable with complication level θ be X_θ. Also, let the corresponding entropy level and the cost be $H(X_\theta)$ and $C(\theta)$. Then, given θ, an individual chooses to approach uncertainty if $V_A(\theta) \equiv V(\theta) - C(\theta) > V_I \equiv 0$, i.e, $V(\theta) > C(\theta)$ and otherwise avoids it, where $V(\theta) = U(X_\theta) = -H(X_\theta)$.

Getting Assisted in Resolving Uncertainty An individual may prefer receiving assistance from others rather than resolving uncertainty for himself. Although an individual may reduce the effort to resolve uncertainty if he gets assisted, his utility is in general reduced as well than if he resolves it for himself. Let the utility when he is helped by others be $V_O \equiv \hat{V} - \hat{C}$ where $\hat{V} < V$ and $\hat{C} < C$. Then, he addresses uncertainty if $\max\{V_A, V_O\} > 0 = V_I$ and resolves it by being helped especially if $V_A < V_O$. Otherwise, he ignores uncertainty.

Partial Resolution Let the effort level be e. We can assume that a higher level of effort gives an individual a more informative signal $Y(e)$ for the uncertainty, implying that $H(X_\theta|Y(e))$ is strictly decreasing. Let $V(e,\theta) = -H(X_\theta|Y(e))$. Then, V increasing in e. We also assume that $C_1(e,\theta) > 0$ and $C_{11}(e,\theta) > 0$. Then, the decision-maker will choose e such that

$$\max_e V(e,\theta) - C(e,\theta).$$

Assuming that $V_{11} < 0$, we have

$$V_1(e^*,\theta) = C_1(e^*,\theta). \tag{4}$$

That is, he will choose the effort level the balance of the incremental change in the benefit and the cost. Differentiating (4) gives the partial derivative

$$de^*/d\theta = (V_{12} - C_{12})/(C_{11} - V_{11}).$$

Note that e^* has the extreme value when $V_{12} = C_{12}$. If V_{12} is decreasing and C_{12} is increasing in θ, it suggests an inverted U-shaped curve which has been observed by psychologists. [19]

8 Conclusion

In this paper, we defined curiosity in terms of information entropy by Shannon and empirically showed that people have curiosity in the sense that they get direct utility from

[19] For the evidence of U-shaped curves, see Hebb (1955) and, Menon and Soman (2002).

information contrary to the expected utility theory. We also interpreted Shannon constant as a degree of personal curiosity and proposed a measure of individual curiosity which is the individual's willingness-to-pay to obtain information. This measure is methodologically distinguished from various measures developed in psychology which rely on pencil-and-paper questionnaires. Moreover, we show that this method of measuring the individual curiosity level in terms of his willingness-to-pay is more informative than particularly IPI curiosity measure used in psychology in the sense that the distribution is more disperse. However, we think that our measure is still not completely satisfactory in the sense that the concept implicitly assumes that all individuals have the same utility from money. If it is not the case, we cannot say that an individual who is willing to pay 100 won for some information is more curious than another who is willing to pay 50 won. Thus, in this case, WTP cannot be a proper measure of curiosity and we need more refined method. One possible alternative method without resorting to some value intermediary such as money would be to use fMRI to compare the image reflecting the individual's utility level directly. We will take up this approach in a different paper.

Although all the concepts offered in this paper are rudimentary and need to be further refined, we believe that the overall perspective should deserve attention and that it can be fruitfully applied to many economic situations. We look forward to future research along this line.

Appendix

Proof of Corollary 1:

Lemma 1 (Jensen's Inequality) *Suppose* $a_i > 0$ *for* $i = 1, \ldots, n$ *and* $\sum_{i=1}^{n} a_i = 1$. *Then, if* $x_i > 0$ *for* $i = 1, \ldots, n$, *we have*

$$\sum_{i=1}^{n} a_i \log x_i \leq \log \sum_{i=1}^{n} a_i x_i$$

with equality if and only if $x_1 = x_2 = \ldots = x_n$.

We have

$$H(X) \quad = \quad -k \sum_{i=1}^{n} p_i \log p_i$$

$$
\begin{aligned}
&= k\sum_{i=1}^{n} p_i \log\frac{1}{p_i} \\
&= \log k\sum_{i=1}^{n}\left(p_i\times\frac{1}{p_i}\right) \quad \text{by Lemma 1} \\
&= k\log n.
\end{aligned}
$$

Proof of Theorem 2: Let us start by defining the joint entropy of X and Y as

$$H(X,Y) = -k\sum_{i=1}^{n}\sum_{j=1}^{m} P(X=x_i, Y=y_j)\log P(X=x_i,Y=y_j),$$

where $P(X=x_i, Y=y_j)$ is the joint probability. Then, we have

Lemma 2 $H(X,Y) = H(Y) + H(X \mid Y)$.

Proof. We have

$$
\begin{aligned}
H(X,Y) &= -k\sum_{i=1}^{n}\sum_{j=1}^{m} P(Y=y_j)P(X=x_i\mid Y=y_j)\log P(Y=y_j)P(X=x_i\mid Y=y_j) \\
&= -k\sum_{j=1}^{m} P(Y=y_j)\left(\sum_{i=1}^{n} P(X=x_i\mid Y=y_j)\right)\log P(Y=y_j) \\
&\qquad -k\sum_{i=1}^{n}\sum_{j=1}^{m} P(Y=y_j)P(X=x_i\mid Y=y_j)\log P(X=x_i\mid Y=y_j) \\
&= H(Y) + H(X\mid Y).
\end{aligned}
$$

Lemma 3 $H(X,Y) \le H(X) + H(Y)$ *with equality if and only if* X *and* Y *are independent.*

Proof: Note that $\log z \le z-1$. Consider $z = \frac{p_i p_j}{p_{ij}}$ where $p_i = P(X=x_i)$, $p_j = P(Y=y_j)$ and $p_{ij} = P(X=x_i, Y=y_j)$. We have

$$
\begin{aligned}
\nabla &= \sum_j\sum_i p_{ij}\log\frac{p_ip_j}{p_{ij}} \\
&\le \sum_j\sum_i p_{ij}\left(\frac{p_ip_j}{p_{ij}} - 1\right) \\
&= \sum_j\sum_i p_{ij} - \sum_j\sum_i p_ip_j \\
&= 0.
\end{aligned}
$$

This implies that

$$H(X,Y) = -k \sum_j \sum_i p_{ij} \log p_{ij} \geq -k \sum_j \sum_i p_{ij} \log p_i p_j = H(X) + H(Y).$$

The proof of Theorem 2 follows immediately from Lemma 2 and Lemma 3.

Proof of Theorem 3: Note that $\phi(p) = p \log p$ is is strictly convex. Hence, the proof is immediate from Blackwell's Sufficiency Theorem. See Blackwell (1951) for Sufficiency Theorem.

References

[1] Arrow, K., 1972, The Amount and Demand for Information, In McGuire C. B. and R. Radner (ed.), *Decision and Organization: A Volume in Honor of Jacob Marschak*, (North-Holland, Armsterdam).

[2] Bellman, R.E. and R. Kalaba, 1957, Dynamic Programming and Statistical Communication Theory, Proceedings of the National Academy of Sciences 43, 749-751.

[3] Blackwell, D., 1951, The Comparison of Experiments, In Neyman J.(ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 93-102.

[4] Byun, D.-K., 2003, South Korea: Women Smokers Start Habit Out of Curiosity, In *The Korea Times*, August 29.

[5] Condry, J., 1977, Enemies of Exploration: Self-Initiated versus Other-Initiated Learning, Journal of Personality and Social Psychology 35, 459-477.

[6] Fidler, A., van Oers K., Drent, P.J., Kuhn, S., Mueller, J.C. and Kempenaers, B., 2007, Drd4 Gene Polymorphisms are Associated with Personality Variation in a Passerin Bird, Proceedings of the Royal Society London B 274, 1685-1691.

[7] Fischbacher, U., 2007, z-Tree: Zurich Toolbox for Ready-Made Economic Experiments: Experimental Economics 10, 171-178

[8] Green, D., 1985, The Phenomenon of Drug Abuse: Its Characteristics in Adolescence and the Role of the Family in Coping with the Problem. In Ziv A. (ed.), *The Unusual Age*, (Tel Aviv: Papyrus). 258-304.

[9] Hebb, D.O., 1955, Drives and the C.N.S. (Conceptual Nervous System), Psychological Review 62, 243-254.

[10] Hirshleifer, J., 1971, The Private and Social Value of Information and Reward to Inventive Activity, American Economic Review 61, 561-574.

[11] James, W., 1950, Principles of Psychology, Vol.2, New York: Holt [Originally published 1890]

[12] Kelly, J.L. Jr., 1956, A New Interpretation of Information Rate, Bell System Technical Journal 35, 917-926.

[13] Kreitler, S., Zigler, E. and Kreitler, H., 1974, The Complexity of Complexity, Human Development 17, 54-73.

[14] Krugman, H., 1965, The Impact of Television Advertising: Learning without Involvement, Public Opinion Quarterly 29, 349-356.

[15] La Valle, I.H., 1968, On Cash Equivalents and Information Evaluation in Decisions under Uncertainty, Parts I, II and III, Journal of the American Statistical Association 63, 252-290.

[16] Litman, J., 2005, Curiosity and the Pleasures of Learning: Wanting and Liking New Information, Cognition and Emotion 19, 793-814.

[17] Litman, J. and Spielberger, C.D., 2003, Measuring Epistemic Curiosity and its Deversive and Specific Components, Journal of Personality Assessment 80, 75-86.

[18] Loewenstein, G., 1994, The Psychology of Curiosity: A Review and Reinterpretation, Psychological Bulletin 116, 75-98.

[19] Lumsdaine, A.A., and Janis, I.L., 1953, Resistance to 'Counterpropaganda' Produced by One-Sided and Two-Sided 'Propaganda' Presentations, Public Opinion Quarterly 17, 311-318.

[20] Marschak, J., 1959, Remarks on the Economics of Information, Contributions to Scientific Research in Management, Western Data Processing Center, University of California, Los Angeles, 79-98.

[21] Marschak, J. and Miyasawa, K., 1968, Economic Comparability of Information Systems, International Economic Review 9, 137-174.

[22] McAlister, A., Perry, C., Killen, J., Slinkard, L.A. and Maccoby N., 1980, Pilot Study of Smoking, Alcohol and Drug Abuse Prevention, American Journal of Public Health 70, 719-721.

[23] Menon, S. and Soman, D., 2002, Managing the Power of Curiosity for Effective Web Advertising Strategies, Journal of Advertising 31, 1-14.

[24] Mizner, G.L., Barter, J.T., and Werme, P.H., 1970, Patterns of Drug Use among College Students: A Preliminary Report, American Journal of Psychiatry 127, 15-2.

[25] Murota, T., 1988, Demand and Supply Values of Information, Information Economics and Policy 3, 25-34.

[26] Naylor, F.D., 1981, A State-Trait Curiosity Inventory, Australian Psychologist 16, 172-183.

[27] Ormian, H., 1975, Adolescence: Portfolio of Chapters for Perusal, Tel Aviv: Otzar Hamore. (Hebrew)

[28] Papageorgis, D. and McGuire, W.J., 1961, The Generality of Immunity to Persuasion Produced by Pre-exposure to Weakened Counterarguments, Journal of Abnormal and Social Psychology 62, 475-481.

[29] Pearson, P.H., 1970, Relationships between Global and Specific Measures of Novelty Seeking, Journal of Consulting and Clinical Psychology 34, 199-204.

[30] Saxe, R.M. and Stollak, G.R., 1971, Curiosity and Parent-Child Relationship, Child Development 45, 892-907.

[31] Schrödinger, E., 1944, What is Life? - The Physical Aspect of the Living Cell, Cambridge University Press: U.K.

[32] Shannon, C.E., 1948, A Mathematical Theory of Communication, The Bell System Technical Journal 27, 379-425.

[33] Singer, J.L. and Antrobus, J.S., 1970, Manual for the Imaginal Processes Inventory, Educational Testing Service: Princeton, New Jersey.

[34] Vidler D.C. and Rawan, H.R., 1983, Curiosity and Exploration: Theories and Results, Academic Press: New York.

[35] Wohlwill, J.F., 1981, A Conceptual Analysis of Exploratory Behavior, In *Advances in Intrinsic Motivation and Aesthetics*, (New York: Plenum Press).

[36] Zuckerman, M., 1979, Sensation Seeking: Beyond the Optimal Level of Arousal, Cambridge University Press: New York.

[37] —, 1994, Behavioral Expressions and Biosocial Bases of Sensation Seeking, Erlbaum: New York.