

정보의 주요 특성이 직관적 확률판정에서의 과신현상에 미치는 영향에 관한 연구

조 성 구*

A Study on the Effects of Information Characteristics on the Overconfidence Phenomenon in Intuitive Probability Judgements

Sung-Ku Cho*

Abstract

Previous studies have shown strong tendency toward overconfidence in intuitive probability judgements. The purpose of the present study is to investigate the relations between this overconfidence phenomenon and the three major characteristics of information, namely, the pertinance, the redundancy and the quantity.

An experiment was conducted where the subjects were asked to respond to 120 questions of the same type. In each question, the subjects' task was to predict, in the light of given information, which of the two given countries would have had higher GNP in 1979 and to give the probability that their choice would be correct.

The results suggests that only the pertinance of information has significant influence on the degree of overconfidence.

1. 序 論

기대효용의 최대화를 의사결정 기준으로 하는 대부분의 의사결정 모델이나 위험을 내포한 일

상의 의사결정 문제들은 불확실한 변수에 대한 확률분포의 추정을 기본적인 선결조건으로 하고 있다. 그러나 실제 상황에서 완벽하게 객관적인 확률 추정치를 얻을수 있는 경우는 거의 없으며,

다소간의 사결정자의 주관적인 요소가 확률판정에 가미되기 마련이고, 또한 유사한 상황에 대한 과거의 데이터가 충분치 못하다거나 신속한 의사결정이 요구되는 상황에서는 전적으로 의사결정자나 전문가의 주관적이고도 직관적인 판단에 의지할 수 밖에 없다.

한편 주관 확률론의 입장에서는 확률은 각 개인이 어떤 미지의 상황에 대해 갖고 있는 불확실성(또는 확신)의 정도를 수치화 한 것에 불과하므로 주관적으로 판정된 확률값들 간에 최소한의 일관성만 있으면 적당한 확률로 받아들여야 한다는 견해를 보이고 있다(DeFinetti, 1975, Howard, 1980). 여기서 말하는 최소한의 일관성이란 규범적 일관성을 가리키는 것으로 확률계산의 기본법칙들을 위배해서는 안된다는 것을 의미한다. 그러나 주관적 확률을 실제 의사결정 문제에 도입하기 위해서는 또다른 측면의 일관성이 요구된다. 즉, 의사결정자가 불확실한 상황에서 어떤 사상의 발생확률을 p 라고 판정하였다면, 실제로 그와 같은 사상의 발생빈도가 p 가 되어야 한다는 것이다.

앞에서 언급한 두가지 일관성중 전자는 비교적 그 검증이 용이하다고 볼수 있는 반면, 후자는 실제적으로 검증이 이루어질 수 있는 경우가 많지 않다. 그러나 서로 유사한 상황에서 많은 수의 확률판정이 반복적으로 이루어지는 경우에는 간접적으로나마 그 검증이 가능한 경우가 있다. 예를 들어 어떤 기상전문가가 오랜 기간에 걸쳐 매일 다음날의 기상상황을 예측하고 그 예측에 대한 확신의 정도를 확률로 나타내 준다면, 이중 이를테면 $x\%$ 의 확률치를 준 날들 중에서 실제로 예측한 기상상황이 발생한 날이 $x\%$ 가 되는가를 확인해 볼수 있다. 다시 말해 다음날 비를 확률을 70%라고 예측한 경우가 총 100회였는데 이 중에서 실제로 70일은 비가 왔다면, 이 기상전문가는 자신이 가지고 있는 불확실성의 정도를 확률로 매우 잘 표시했다는 것을 알

수 있으며, 이때 이 기상전문가의 확률판정은 “잘 교정되어 있다(well-calibrated)”고 말한다. 그러나 만약 실제로 비온 날이 전체의 60%나 또는 80%라고 한다면 이 기상전문가의 확률판정에는 문제가 있다고 말할 수 있다.

여러 다른 상황에서의 확률판정의 교정(calibration) 정도를 알아본 경험적 연구는 많이 있었는데, 이러한 연구에서 나타난 보편적 현상은 대부분의 확률판정자들이 자신의 예측능력을 너무 과신하고 있다는 것이다(Lichtenstein et al., 1982). 위의 예를 다시 든다면, 비를 확률을 70%라고 한 경우 중 실제로 비온 날의 수는 전체의 70%가 훨씬 못되는 경우가 대부분이라는 것이다. 이와 같은 현상을 가리켜 주관적 확률판정에서의 과신현상(phenomenon of overconfidence)이라고 부르는데, 본 연구에서는 이와 같은 현상의 보편성을 다시 확인해 보고 확률판정의 과신정도(degree of overconfidence)에 영향을 미치는 요인은 어떤 것들이 있는지 실험을 통하여 알아보려고 한다. 이러한 요인들로는 여러가지를 생각해 볼수 있겠으나, 여러 선행연구를 통하여 볼 때 과신현상에 대한 영향이 특히 클 것으로 생각되는 정보의 양, 적합성, 중복도(또는 일관성) 등 정보의 제반 특성들에 대하여 가설을 세우고 이를 검증해 보고자 한다.

2. 用語의 定義

가설의 수립 및 실험의 내용과 결과를 소개하기 앞서 본 연구에서 사용할 핵심되는 용어들에 대한 정의를 먼저 정확히 해 둘 필요가 있다.

2-1. 교정(Calibration)의 척도

대략적인 교정의 의미는 서론에서 잠깐 언급하였으나, 여기서는 교정의 척도를 정확히 정의 하도록 한다. 확률판정을 하려는 대상 변수의 특

성에 따라 여러가지 교정의 척도를 사용할 수 있지만, “예-아니오”, “성공-실패” 등 표본공간이 두가지 대립사상으로 분할되는 경우에는 Murphy(1973)가 제시한 교정의 척도가 가장 적합한 것으로 많이 사용된다.

어떤 확률판정자가 “예-아니오” 등 두가지 대답만이 가능한 일련의 유사한 질문에 대하여 답하고 매번 자신의 예측이 맞을 가능성을 확률로 나타내 준다면, 이러한 일련의 확률판정에 대하여 교정의 정도는 다음과 같은 척도로 측정할 수 있다.

$$\text{교정} = \frac{1}{N} \sum_{t=1}^T n_t (r_t - c_t)^2$$

단,

N =응답(즉, 확률판정)의 총 횟수,

T =응답 범주(response category)의 총 수. 여기서 응답 범주란 같은 확률값을 준 응답의 집합을 말한다.

r_t =응답 범주 t 에 부과된 확률,

n_t =범주 t 에 속하는 응답의 총 수. 즉, 확률 값 r_t 가 사용된 총 횟수.

따라서, $\sum_{t=1}^T n_t = N$ 이다.

c_t =범주 t 에 속한 응답의 정확도, 즉, v_t 는 범주 t 에 속한 응답중 정확한 답의 총 수를 말한다면 $c_t = v_t/n_t$ 를 의미한다.

따라서 교정의 척도는 각 범주에 속한 응답의 정확도와 그 범주에 속한 응답들에 부과된 확률과의 차이를 자승하여 응답수에 따라 가중 평균한 값이다. 완벽히 교정된 확률판정자라면 이 척도로 0값을 받게 될 것이고, 반대로 모두 틀린 답을 제시하였으면서 정확히 맞출 확률은 모두 1을 부과한 경우에는 교정의 값이 1이 될 것이다.

1-2. 과신(Overconfidence)의 척도

일반적으로 확률판정에 있어서의 과신현상은

완벽하게 교정이 되었을 경우에 얻게될 확률분포에 비하여 지나치게 “뽀족한” 분포를 주는 경향을 일컫는 것이다. 그러나 앞에서 정의했던 교정의 척도에서와 마찬가지로, 두가지 대답만이 가능한 일련의 질문에 대하여 답하고 그 답이 맞을 확률을 제시하는 경우에는 다음과 같은 과신의 정도에 대한 척도를 사용할 수 있다(Lichtenstein & Fischhoff, 1977).

$$\begin{aligned} \text{과신의 정도} &= \frac{1}{N} \sum_{t=1}^T n_t (r_t - c_t) \\ &= r - c \end{aligned}$$

단,

$$r = \frac{1}{N} \sum_{t=1}^T n_t r_t$$

=모든 응답에서 부과된 확률의 평균.

$$c = \frac{1}{N} \sum_{t=1}^T n_t c_t$$

=전체적인 응답의 정확도.

따라서 위와 같은 척도를 일련의 확률판정에 적용하였을 때, 양의 값을 얻는다면 과신(overconfidence)을 나타내는 것이고 음의 값을 얻는다면 자신감 결여(under-confidence)를 나타내는 것으로 볼수 있다.

일련의 확률판정에 대해 확률판정자가 부과한 확률과 응답의 정확도간의 관계를 나타낸 곡선을 가리켜 교정곡선(calibration curve)이라 하는데 다음의 그림 1은 교정곡선의 몇가지 예를 보여준다.

여기서 직선 B는 완벽한 교정을 나타내고, 곡선 A는 자신감 결여, C는 과신의 경우 나타나는 전형적인 교정곡선이다. 예를 들어 교정곡선 C를 보면 80%의 확률을 부과한 응답중 60%만이 정답이었음을 알수 있다.

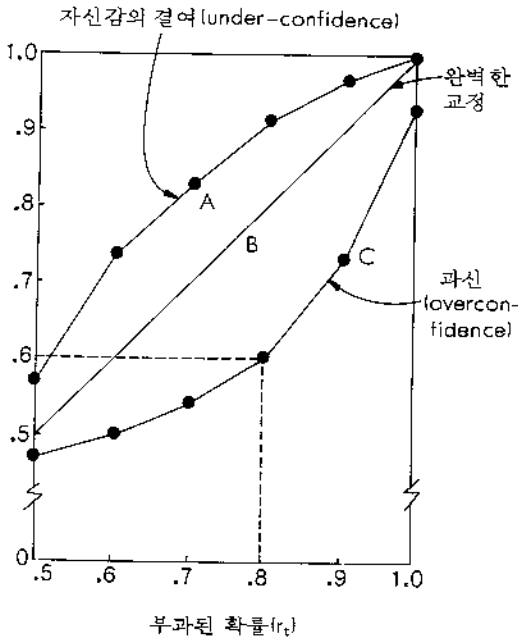


그림 1. 교정곡선의 예

3. 정보의 제반특성과 과신현상

교정이나 과신현상에 대한 많은 선행연구들 중 정보의 여러 특성이 과신현상에 미치는 영향을 조사한 연구는 그다지 많지 않다. 그러나 연구의 목적이 정보의 특성과 직접 관련된 것은 아니라 하더라도 연구의 결과들을 비교 검토하면 정보의 몇 가지 중요한 특성이 과신의 정도에 상당한 영향을 줄 것으로 짐작되는 경우가 많다.

본 연구에서는 특히 정보의 내용이 확률판정을 하려는 변수와 얼마나 연관이 있는가를 나타내는 정보의 적합성(pertinence), 정보간에 중복된 정도를 나타내는 정보의 중복도(redundancy) 그리고 정보의 양(quantity)에 대하여 이와 같은 관계를 조사하여 보고자 한다.

3-1. 정보의 적합성과 과신현상

정보의 내용이 예측하고자 하는 변수와 높은 상관관계를 갖고 있을 때와 그렇지 못할 때에 따라 이 정보를 이용하여 내려진 확률판정의 교정 정도는 어떻게 달라질 것인가를 직접적으로 조사한 연구는 별로 없다. 그러나 여러 선행연구들을 살펴 볼 때 일반적으로 직관적인 확률판정은 정보의 가치에 매우 둔감하다는 것을 알 수 있다. 예를 들어 직관적인 확률판정은 표본의 크기나 (Kahneman & Tversky, 1972) 정보의 신뢰성 (Hogarth, 1980, p. 33; Schum et al., 1973) 등에 별로 영향을 받지 않으며, 그 결과 표본의 크기가 크거나 정보의 신뢰도가 높으면 정보의 가치를 충분히 반영하지 못하는 현상이 일어나고, 반대로 표본의 크기가 매우 작거나 정보의 신뢰도가 낮으면 정보의 가치를 과대평가하여 과신하는 현상이 발생한다.

이러한 연구결과를 비추어 볼 때, 정보의 적합성에 대해서도 유사한 결과를 얻을 수 있을 것으로 예측하여 볼 수 있다. 즉, 정보의 적합성이 낮으면 낮을수록 과신의 정도는 더 심해질 것으로 예측된다. 따라서 본 연구에서는 예측하려는 변수와 정보로 주어진 변수와의 선형 상관계수를 정보의 적합성을 나타내는 척도로 사용하여 다음과 같은 가설을 세우고 검증하여 본다.

가설 1: 직관적인 확률판정은 정보의 적합성을 충분히 고려하지 못하므로, 정보의 적합성이 낮을수록 과신의 정도는 더 커진다.

3-2. 정보의 중복도와 과신현상

정보의 중복도나 일관성이 확률판정에 미치는 영향에 관하여 Kahneman 과 Tversky(1973)는 몇 가지 실험을 해 본 결과 다음과 같은 현상을 발견하였다. 다수의 정보를 가지고 예측을 하는 경우, 예측에 대한 확신의 정도는 주어진 정

보들의 일관성이 높을수록 커지지만, 일반적으로 일관성이 높은 정보들은 서로간의 상관관계가 높은 것이 대부분이므로, 각 정보의 가치가 같다고 할 경우 오히려 전체적인 정보의 가치는 정보들간의 중복도가 클수록 줄어들게 되어 정보들간의 상관관계가 높을수록 예측에 대한 확신은 커지는 반면, 예측의 정확도는 떨어지는 역설적인 현상이 발생한다는 것이다. 바꾸어 말하면, 이것은 정보의 중복도가 높을수록 과신의 정도가 심해진다는 것을 의미한다.

한편, 이와 같은 사실을 확인하기 위하여 Kahneman 과 Tversky 가 행한 실험은 어떤 과목에서 얻은 학점 등과 같이 최소한 순위가 정해져 있는(ordinal scale) 변수를 예측할 변수로 사용하였는데, 이 경우 그들의 결론은 매우 타당성이 있는 것으로 생각된다. 그러나 몇가지 병 중에서 어떤 증상의 원인이 되는 병은 무엇인가를 판정한다든지 또는 대학입시에서의 합격, 불합격을 판정하는 경우에서와 같이 예측하려는 변수가 nominal scale 로 되어 있는 경우에도 Kahneman 과 Tversky 의 결론이 그대로 적용될 수 있을지는 전혀 미지수다.

따라서 본 연구에서는 주어진 정보들간의 선형 상관계수를 정보의 중복도의 척도로 사용하여 다음과 같은 가설을 세우고 검정한다.

가설 2: 다수의 정보에 근거하여 확률판정이 이루어지는 경우, 정보들간의 상관관계가 높을수록 과신의 정도가 커진다.

3-3. 정보의 양과 과신현상

많은 수의 정보가 주어졌을 경우 정보들간의 중복도 뿐만 아니라 정보의 양 자체도 확률판정에 영향을 줄 것으로 생각될 수 있다. 정보처리 능력에 한계가 없다면, 이론적으로 더 많은 정보를 얻었을 경우 예측의 정확도와 예측에 대한 확신의 정도는 모두 높아지게 될 것이다. 그러나

인간의 정보처리 능력에는 한계가 있고, 처리방식도 최적이라고 할수는 없으므로 정보량의 증가가 과신의 정도에 어떤 영향을 미칠 것인가는 단적으로 말하기 어렵다.

이 문제와 관련하여 마케팅 분야에서는 이미 흥미있는 연구결과가 알려져 있다(Jacoby et al., 1974). 제품에 관한 정보의 양을 늘려가면 소비자들의 판단의 정확도도 상승하나 어느 수준 이상의 정보량은 오히려 판단의 정확도를 떨어뜨린다는 것이 그것이다. 그렇다면 판단에 대한 확신도 정확도에 따라 비례하여 변할 것인가?

Oskamp(1965)는 정신과 의사들을 대상으로 이러한 사실을 검증해 보려 하였다. 그는 몇가지 정신질환 케이스를 이용하여 실험해 본 결과 4가지의 정보를 하나씩 추가하여 줄 때마다 피실험자들의 확신은 높아지는데 반하여 정확도에 상응하는 향상이 없었으므로, 정보의 양이 증가될수록 과신의 정도만 높아진다고 결론짓고 있다.

한편 Oskamp 의 실험을 일반화 시키는에는 몇가지 문제점이 있는데, 첫째는 그가 사용한 문제들이 너무 어려웠다는 것이고(무작위로 대답한 경우보다 정확도가 별로 높지 않았음), 둘째는 현실문제에서와 같이 4가지 이상의 더 많은 정보를 주는 경우는 어떻게 될 것인가 하는 점이다.

따라서 본 연구에서는 문제의 종류와 정보의 양을 달리하여 Oskamp 의 결론을 재 확인해 보고자 한다.

가설 3: 정보의 양이 증가할수록, 예측에 대한 확신이 증가하는 만큼 예측에 대한 정확도는 증가하지 않으므로, 과신의 정도는 커진다.

4. 실험의 方法과 절차

앞에서 제시한 3가지 가설을 검정해 보기 위

하여 본 연구에서는 실험 심리학적 방법을 연구 방법으로 채택하였다. 특히 과신현상에 관한 연구라는 점을 고려하여 피실험자들에게 일련의 확률추정 문제를 제시하고 응답하도록 하되 가급적이면 i) 문제의 성질이 모두 유사하고 반복적인 것이며, ii) 피실험자는 주어진 정보만을 사용하여 관정하되 정보의 제반특성은 실험자가 임의로 조정할 수 있는 것으로 하였고, iii) 주어진 문제가 너무 가상적인 것이 아니고, 피실험자에게 현실적인 의미를 갖는 것으로 선정하도록 하였다.

따라서 본 연구에서는 많은 수의 확률추정 문제를 만든 다음 이것을 설문지의 형태로 정리한 후, 피실험자들에게 제시하여 일정시간 동안 응답하도록 하고, 그 결과를 분석하여 앞에서 제시한 가설들을 검정해 보는 절차를 따랐다.

4-1. 피실험자

실험에 참가한 피실험자는 총 60명이었고, 모두 대학 2년 이상의 교육을 받은 사람들이었다. 평균 나이는 34세이고, 전체의 절반 가량이 30세에서 40세 사이였으며, 최연소자는 26세, 최연장자는 58세 였다. 피실험자들의 직업분포는 다음과 같다.

엔지니어	기업체 중간간부	금융관계 전문가	총계
28	26	6	60

4-2. 실험절차

실험은 몇회로 나누어 피실험자를 한자리에 모이게 한 후 설문지를 나누어 주고, 약 10분간 설문지 응답방법, 교정의 개념, 상관계수의 개념 등을 설명하고, 피실험자가 원하는 속도로 응

답하도록 하는 절차를 따랐다. 시간의 제약은 두지 않았으나 전체적으로 2시간을 넘긴 피실험자는 없었으며, 평균 응답시간은 65분이다. 설문지 자체의 단조로움을 호소한 경우가 많았으나 전체적으로는 협조적 분위기였다.

4-3. 설문지의 구성

본 연구에서 설문지의 구성은 매우 중요하다. 학생들을 대상으로 한 1회의 예비실험을 거친 후 설문지의 최종 형태를 결정하였는데, 설문지는 모두 3부분으로 구성되도록 하였다.

1) 제 1부

제 1부는 30개의 상식문제로 되어 있는데, 어느 것이나 두가지 가능한 대답중 하나가 반드시 정답이다. 피실험자는 이 중 정답으로 생각하는 하나를 선택한 후 자신의 답이 맞았을 확률을 %로 표시하도록 되어 있다. 한 예를 들면 다음과 같다.

예 : 국제 사면기구(Amnesty International)의 창설자는

Ⓐ. 영국인이다.

B. 미국인이다. 70 %

제 1부의 내용이 본 연구의 목적과 직접 관련이 있는 것은 아니지만 과신현상에 관한 많은 선행연구가 이와 같은 상식문제를 실험도구로 사용한 예가 많기 때문에 그 결과와 비교하여 봄으로써 본 연구의 피실험자들이 예외적인 성향을 가지지는 않았는가 확인하기 위한 것이 제 1부의 목적이다. 전반적으로 질문의 내용은 평이한 것과 어려운 것 등이 골고루 배분되도록 하였다.

2) 제 2부

제 2부는 본 연구에서 제시한 가설들을 검정

하기 위한 핵심부분으로서, 모두 120개의 문항으로 구성되어 있다. 각 문항은 제시된 정보의 숫자만 다를 뿐 모두 주어진 경제사회 데이터를 이용하여 익명으로 되어 있는 두 나라중 어느 나라의 GNP가 1979년에 더 높았을 것인가를 예측하고, 그 예측이 맞을 확률을 피실험자가 주목 되어 있다. 한 예를 들면 다음과 같다.

예 :	국가 A	국가 B
총수출액 (단위 : 10억\$)	2.07	15.58
인구 (단위 : 萬명)	815	14,847
	95 %

이와 같은 문제유형은 정답을 곧 알수 있고, 유사한 문제를 많이 만들어 낼수 있을 뿐만 아니라 여러 통계자료를 통해 나와 있는 각국의 경제사회 데이터를 이용하면 변수들의 상관관계를 쉽게 알아낼 수 있는 이점이 있다. 또한 두 국가의 이름이 익명으로 되어 있으므로 피실험자는 오직 주어진 정보에만 의존해서 판정을 할수 밖에 없다.

한편 120개의 문항을 만들기 위하여 통계연감등을 통해 데이터를 구할 수 있는 125개국의 GNP를 위시한 인구, 면적, 에너지 소비량, 전기생산량, 실업률, 교육수준 등 총 21개의 경제사회적 변수들간의 상관관계를 SPSS를 이용하여 컴퓨터로 분석하였다. 사용된 변수들의 자세한 목록과 계산된 상관관계 행렬의 일부는 제 5절의 표 3을 참조하기 바란다(GNP를 제외한 변수들간의 상관관계는 생략한다).

이와 같이 얻은 데이터를 기초로 하여 만들어진 120개의 문항은 주어진 정보의 수에 따라 4군으로 나누어 설문지에 수록하였다.

첫번째 문제군(총 40문항)은 본 연구의 가설 1을 검증하기 위한 것으로 각 문항당 한가지씩의 정보만이 주어져 있는데, 그중 20 문항은 GNP와의 상관관계가 높은 변수의 값이 정보로 주어

지고, 나머지 20 문항은 GNP와의 상관관계가 낮은 변수의 값이 정보로 주어져 있다. 또한 두번째군 40 문항은 가설 2를 검증하기 위한 것으로 각 문항당 2개씩의 정보가 주어져 있다.

그룹 1에서와 마찬가지로 이중 20 문항의 경우는 정보로 주어진 두 변수간의 상관관계가 매우 높고, 나머지 20 문항은 그 반대의 경우다. 끝으로 세번째와 네번째 군은 각각 주어진 정보의 수가 4와 10이고, 각 그룹당 20 문항씩으로 모두 가설 3을 검증하기 위하여 만들어진 것이다.

이와같이 하여 120개의 문항을 만드는 과정에서 조사해 보고자 하는 정보의 특성 이외의 다른 요인이 개입되지 않도록 각 변수의 이용도, 예상되는 확률값의 배분 등이 골고루 이루어지도록 세심한 주의를 기울였고, 문제의 순서도 가급적 랜덤이 되도록 하였으나 그 자세한 과정은 여기서 생략하기로 한다.

3) 제 3부

제 3부는 2부의 부속편으로, 정보로 사용되었던 20개의 경제사회 변수가 각각 GNP와 어떤 상관관계를 갖는다고 생각되는지 피실험자의 의견을 쓰도록 되어 있다. 이것은 2부의 실험결과를 분석하는데 보조자료로 사용하기 위한 것으로, 피실험자의 응답의 일관성 여부를 검토해 본 다거나 확률판정 과정이 어떤 내적인지 체계를 통해 이루어지는지를 간접적으로나마 확인해 보기 위한 자료를 얻기 위함이다.

5. 실험결과의 분석

5-1. 분석방법

피실험자가 위에서 설명한 설문지에 모두 응답하자면 제 3부를 제외하더라도 총 150번의 확

틀판정(제 1부 30 문항과 제 2부 120 문항)을 하여야 한다. 특히 제 2부의 120 문항은 위에서 설명한 바와 같이 각 20 문항씩 모두 6가지의 실험 조건을 적용한 결과다.

따라서 분석의 절차는 우선 각 응답자별로 실험조건에 따라 앞 2절에서 정의한 과신의 정도를 계산하고, 이를 응답자 전체에 대해 규합하여 전체적인 과신의 정도와 교정곡선 등을 얻은 후 필요한 통계적 검정을 통해 가설을 검정하는 절차를 따르게 된다.

5-2. 상식문제에 대한 실험결과

표 1과 그림 2는 설문지 제 1부의 상식문제 30 문항에 대한 전체 피실험자들의 응답 내용을 요약한 것이다. 표 1을 보면 각 문항에 대해 부과된 확률값을 6개의 범주로 나누었는데, 확률 100%는 그 의미나 빈도수로 보아 별도의 한 범주로 취급하기로 하였다(응답의 범위를 50~100%만으로 한 것은 가능한 대답이 언제나 두가지 뿐이었고, 피실험자는 그중 정답으로 생각하는 답을 골랐기 때문임).

표의 내용은 예를 들어, 부과된 확률값이 70~79% 사이인 것은 모두 70%인 것으로 가정하였을 때($r_i = 0.7$), 이 범주에 속하는 확률판정의 횟수는 전 피실험자를 통하여 모두 188회였으며, 이중 정답을 맞춘 빈도수는 100회이고, 따라서 이 범주의 응답 정확도는 53%가 됨을 나타내고 있다. 또한 표를 보면 전반적인 응답의 정확도는 55%이고 과신의 정도는 0.27로 매우 높음을 알수 있다.

그림 2는 표 1을 이용하여 교정곡선을 그려본 것이다. 일견하여 피실험자들의 교정 정도는 극히 불량하며, 심한 과신현상을 나타내고 있다.

이러한 결과는 상식문제를 사용한 선행연구(Lichtenstein & Fischhoff, 1977, 1980; Fischhoff et al., 1977)의 결과보다 더 극심한 것이

표 1. 상식문제에서의 과신정도를 계산하기 위한 표

과신정도=0.27

범주(%)	r_i (%)	n_i	V_i	c_i (%)
50 - 59	50	213	106.5*	50
60 - 69	60	232	106	46
70 - 79	70	188	100	53
80 - 89	80	235	113	48
90 - 99	90	154	90	58
100	100	776	482	62
	-	1798**	997.5	55

* 확률 50%의 의미를 최대한 살리기 위하여, 범주 50-59에 대해서는, v_i 의 값은 $v_i = n(0.5 < p < 0.6) + (1/2)n(p=0.5)$ 으로 계산하였다.

여기서 $n(p=0.5)$ 는 확률 50%를 부과한 응답의 수를 의미하고, $n(0.5 < p < 0.6) = n_i - n(p=0.5)$ 를 의미한다.

** N=1800이 아닌 이유는 일부 응답되지 않은 문항이 있기 때문이다.

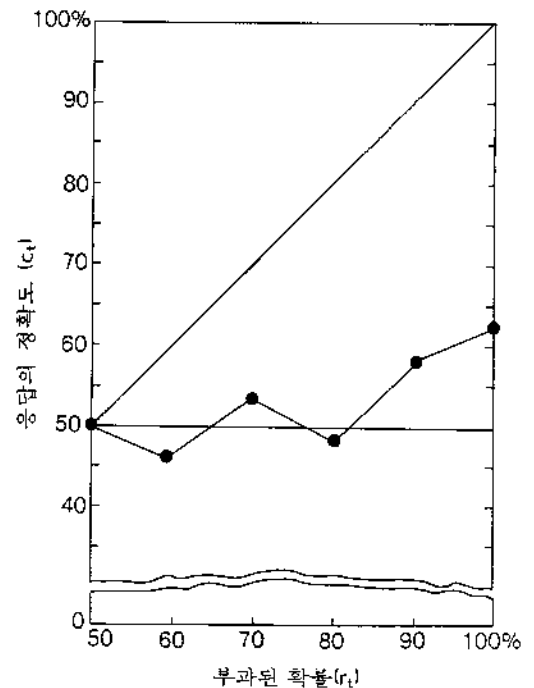


그림 2. 상식문제에 대한 교정곡선

다. 그러나 이러한 현상의 이유가 본 연구의 피 실험자들이 특별히 교정이 불량하기 때문이라기 보다는, 의도적은 아니었으나 응답의 정확도가 55%밖에 안된다는 사실에서 나타나듯이 전반적으로 문제들이 어려웠고, 함정이 많았던 때문이라고 생각된다. 문제의 난이도가 높을수록 과신 현상이 심해진다는 것은 이미 Fischhoff et al. (1977)에서 지적된 바 있다.

한편 이러한 결과는 과신현상에 관한 연구에 상식문제를 사용하는 실험방법이 매우 취약점이 많은 방법이라는 것을 시사하고 있다 하겠다.

5-3. GNP 문제에 대한 전반적인 실험 결과

설문지 제 2부에 해당되는 GNP 문제 120 문항에 대한 전체 피실험자들의 응답내용은 표 2와 그림 3에 요약되어 있다.

전반적인 과신정도는 0.11로 높은 편이며, 교정곡선을 볼때 이것은 특히 80% 이상의 확률을 부과한 문항에 대해 응답의 정확도가 저조했던 때문임을 알수 있다. 이러한 현상은 대부분의 과신현상 연구에서 공통적으로 나타나는 현상으로, 사람들이 1이나 0에 가까운 확률의 부과에 어려움을 많이 느낀다는 사실을 재확인해 주는 것이라 할수 있다. 특히 100% 확률을 남발하는 현상에 주목할 필요가 있는데, 표 2를 보면 이 경우의 응답 정확도가 69%에 불과하다.

5-4. 설문지 제 3부의 결과

설문지 제 3부에서는 정보로서 주어진 변수와 GNP 간의 상관계수를 얼마로 보고 있는지 묻고 있는데, 이에 대한 응답내용을 살펴보면, 우선 상당수의 피실험자가 상관계수의 개념을 잘 모르고 있는 것으로 판명된다. 짧은 시간이나마 상관계수의 개념을 미리 설명하여 주었음에도 불

표 2. GNP 문제에서의 과신정도를 계산하기 위한 표

과신정도=0.11

범주(%)	r_i (%)	n_i	V_i	c_i (%)
50 - 59	50	870	454	52
60 - 69	60	1319	807	61
70 - 79	70	1235	840	68
80 - 89	80	1477	965	65
90 - 99	90	1065	759	71
100	100	1209	833	69
	-	7175	4658	65

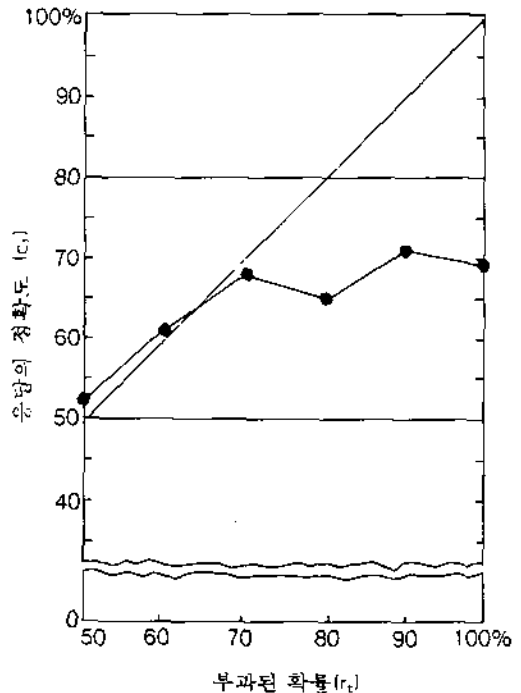


그림 3. GNP 문제 전체에 대한 교정곡선

구하고 제 3부에 응답한 총 57명중 34명이 0, 1, -1이라는 값만을 사용하였다. 따라서 이런 응답은 고려대상에서 제외시키고, 나머지 응답의 내용을 표 3에 요약하였다.

표 3. 피실험자들이 예측한 GNP 와의 상관계수의 평균과 분산

변 수	실제의 상관계수	피실험자가 추정한 상관계수	
		평균	분산
인 구	0.37	0.35	0.54
면 적	0.57	0.19	0.54
단위면적당 목물생산량	0.38	0.47	0.38
에너지 소비량	0.96	0.61	0.29
전기생산량	0.96	0.66	0.21
자동차 생산량	0.91	0.54	0.32
강철생산량	0.89	0.64	0.29
시멘트 생산량	0.78	0.37	0.47
실업률	-0.06	-0.05	0.49
물가상승률	-0.11	0.12	0.52
수출액	0.82	0.47	0.46
수입액	0.86	0.22	0.47
교육수준	0.54	0.53	0.29
광물생산량	0.80	0.44	0.38
신문발행 부수	0.85	0.31	0.34
정부예산의 국방예산 비율	0.02	0.05	0.46
정부예산의 사회보장예산 비율	0.35	0.24	0.41
정부예산의 교육비예산 비율	0.33	0.41	0.41
정부예산의 보건비예산 비율	0.24	0.23	0.51
생산직 근로자의 근로시간	-0.12	0.31	0.41

표 3을 보면, 각 변수가 GNP 와 어떤 상관관계를 가지고 있는지 대체적인 파악은 잘 되어 있다고 볼수 있으나 높은 상관관계를 갖는 변수에 대해서는 그 크기를 과소평가하고 있음을 알수 있다. 특히 주의를 요하는 것은, 전반적으로 분산이 크다는 점으로 알수 있듯이 개인차가 매우 심하다는 사실이다.

5-5. 정보의 적합성과 과신정도

정보의 적합성이 높으면 과신의 정도는 낮아지고, 반대로 적합성이 낮으면 과신의 정도가 심해질 것이라는 것이 가설 1의 내용이였다. 본 연구에서는 GNP 문제를 이용하여 한가지 정보만을 주되 그 정보의 적합성이 높은(즉 GNP 와의

상관관계가 높은) 문항 20개와 적합성이 낮은 경우의 문항 20개를 각각 응답토록 하였는데, 그 결과는 표 4, 5 및 그림 4에 요약되어 있다.

적합성이 높은 정보를 준 경우의 과신정도는 -0.07로서 자신감 결여의 현상이 나타났고, 적합성이 낮은 정보를 준 경우는 과신정도가 0.17이라는 높은 값을 기록했다. 60명의 피실험자 전원이 한명의 예외도 없이 이와같은 현상을 나타냈는데, 이것은 어떠한 통계적 기법을 사용하여 검정하더라도 가설 1이 받아들여질 것이라는 것을 의미한다.

한편 표 4와 5의 n_t 값과 v_t 값을 비교하여 보

면 어디에서 이와같은 명백한 차이가 생겼는가를 짐작할 수 있다. n_t 값은 양쪽의 경우 서로 비슷한 값인 반면에 v_t 값은 표 5의 경우가 훨씬 낮은데, 이것은 피실험자들이 대략 비슷한 방식으로 확률값을 부여하였으나 정보의 적합성의 차이로 인해 응답의 정확도는 많은 차이를 보인 것으로 생각할 수 있다. 이것을 확인해 보기 위하여 n_t 의 분포가 서로 같은지 χ^2 검정을 실시하였으나 통계적으로는 입증할 수 없었다($\chi^2=24.78, p=0.001$).

또 한가지 생각해 볼수 있는 가능성은 피실험자들이 두 종류의 정보들간의 적합성의 차이를

표 4. 정보의 적합성이 높은 경우의 과신정도

과신정도 = -0.07

범주(%)	r_t (%)	n_t	V_t	c_t (%)
50 - 59	50	221	118	53
60 - 69	60	269	212	79
70 - 79	70	176	154	88
80 - 89	80	171	148	87
90 - 99	90	153	146	95
100	100	209	183	88
	-	1199	961	80

표 5. 정보의 적합성이 낮은 경우의 과신정도

과신정도 = 0.17

범주(%)	r_t (%)	n_t	V_t	c_t (%)
50 - 59	50	228	116	51
60 - 69	60	260	138	53
70 - 79	70	218	137	63
80 - 89	80	219	120	55
90 - 99	90	103	54	52
100	100	169	95	57
	-	1197	660	55

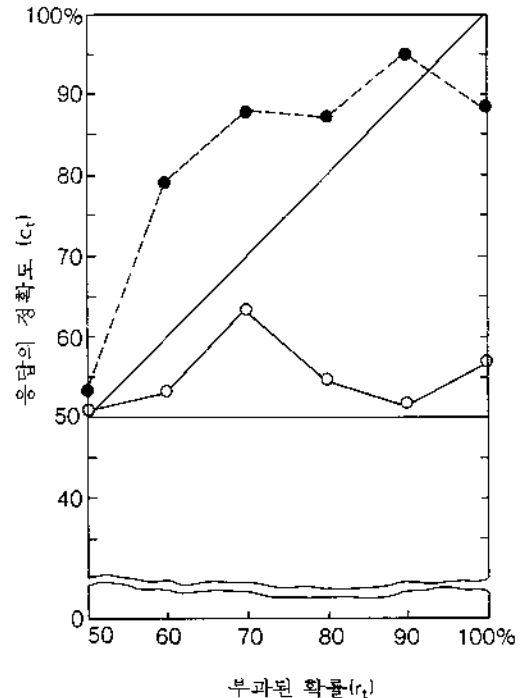


그림 4. 정보의 적합성이 높은 경우와 낮은 경우의 교정곡선

별로 크다고 느끼지 않은 때문이 아닌가 하는 것인데, 표 3의 내용을 살펴보면 이러한 설명만으로 엄청난 과신정도의 차이를 설명하기에는 무리가 따른다고 하겠다.

이상에서 살펴본 내용에 비추어 좀 더 만족스러운 설명은 피실험자들이 어떠한 방식으로 각 문항에 답하고 확률을 부과했는가 하는 것을 생각해 보는 것이다. 정확하게 실험적으로 입증된 것은 아니나 다음과 같은 과정을 통해 확률판정이 이루어진 것이 아닌가 추측된다. 즉,

“피실험자는 주어진 변수가 GNP 와 양, 또는 음의 상관관계를 가졌는가, 혹은 상관관계가 없는가 하는 정도로 대략적으로 판단한 다음, 이에 따라 대답을 선택하고, 확률부과는 정보로 주어진 변수값의 차가 얼마나 큰가로 결정한다.”

만약 이와같은 판정방식이 실제로 사용되었다면 그 결과는 정보의 적합성에 매우 둔감한 확률판정이 될 것이라는 것은 자명하다.

5-6. 정보의 중복도와 과신정도

Kahneman 과 Tversky 는 ordinal scale 로 된 변수의 경우, 정보의 중복도가 높으면(즉, 정보들간의 상관계수가 높으면) 과신현상이 심해진다고 하였는데 nominal scale 로 된 변수의 경우에도 같은 현상이 나타날 것인가를 알아보는 것이 가설 2의 목적이었다. 본 실험에서는 서로 상관관계가 높은 두가지 정보를 제공한 20개의 GNP 문제와 그 반대 경우의 문제 20개에 대하여 피실험자들이 답하게 하였는데, 결과는 표 6과 7, 그리고 그림 5에 요약되어 있다.

실험의 결과를 보면 심하지는 않으나 가설 2와 정반대의 현상이 나타났음을 알 수 있다. 다시 말해 정보의 중복도가 높은 경우의 과신정도는 0.14인데 반하여 중복도가 낮은 경우의 과신정도는 0.20으로 오히려 크다. 특히 그림 5의 교

표 6. 정보의 중복도가 높은 경우의 과신정도

과신정도=0.14

범주(%)	r_t (%)	n_t	V_t	c_t (%)
50 - 59	50	91	47	52
60 - 69	60	200	116	58
70 - 79	70	205	129	63
80 - 89	80	284	195	69
90 - 99	90	223	156	70
100	100	194	123	63
	-	1197	766	64

표 7. 정보의 중복도가 낮은 경우의 과신정도

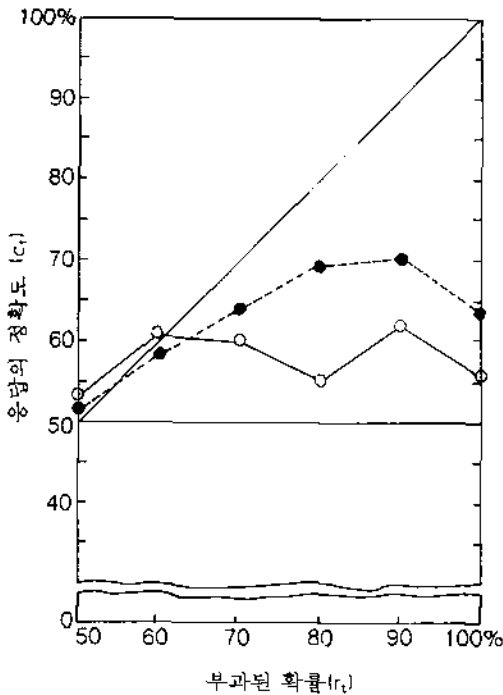
과신정도=0.20

범주(%)	r_t (%)	n_t	V_t	c_t (%)
50 - 59	50	122	64	53
60 - 69	60	178	108	61
70 - 79	70	191	115	60
80 - 89	80	277	155	56
90 - 99	90	216	133	62
100	100	216	121	56
	-	1200	696	58

정곡선을 보면 정보의 중복도가 낮은 경우에는 곡선이 거의 수평을 이루어 피실험자들이 확률판정에 매우 어려움을 겪었다는 것을 알 수 있다.

가설 2와 반대현상이 나타난 정도가 얼마나 심각한지를 알아보기 위해, Wilcoxon 검정을 시행해 본 결과, $T=670$ 을 얻었으며, 정규분포로 근사시켰을 경우 $Z=-1.62$, $p=0.053$ 이 된다. 다시 말해 가설 2를 반대로 설정하였더라면 유의수준 $\alpha=0.10$ 으로 채택되었을 것이다.

한편 표 6과 7을 비교 검토하여 보면, n_t 의 값들이 서로 큰 차이가 없음을 발견할 수 있다. 실



●● : 중복도가 높은 경우
○○ : 중복도가 낮은 경우

그림 5. 정보의 중복도가 높은 경우와 낮은 경우의 교정곡선

제로 n_t 의 분포가 같다고 볼수 있는지 χ^2 검정을 행하여 본 결과 $\chi^2=7.66$, $p>0.10$ 으로 나타났다. 따라서 과신정도의 차이는 주로 응답의 정확도의 차이에서 온 것이라 말할 수 있다.

이상의 결과를 종합하여 볼 때, 가설 2가 기각된 사실은 그다지 놀라운 것이 아니다. 예측하려는 변수가 nominal scale로 된 경우는, ordinal scale의 경우와 달리 정보들간에 일관성이 없다는 것이 그만큼 정보의 추론적 가치가 감소한다는 것을 의미하므로 예측에 대한 확신을 높이지 않는 것은 당연하다. 오히려 정보들간에 중복도가 낮을때는 그만큼 예측이 어려워져 정확도만 떨어뜨리는 결과를 가져온다고 볼수 있다.

5-7. 정보의 양과 과신정도

설문지에는 4가지의 서로 다른 정보가 주어진 문항 20개와 10가지 정보가 주어진 문항 20개가 들어 있다. 그러나 이 20개의 문항들은 서로 같은 것들로 후자의 경우 단지 6가지의 정보만 추가시킨 것이고, 피실험자가 눈치채지 못하도록 문제의 순서와 정보의 제시 순서를 뒤바꾸어 놓은 것이 다를 뿐이다.

가설 3은 정보의 양이 계속 증가될 경우, 응답의 정확도는 별로 증가하지 못하나 응답에 대한

표 8. 4가지 정보를 준 경우의 과신정도

과신정도=0.10

범주(%)	r_t (%)	n_t	V_t	c_t (%)
50 - 59	50	108	57	53
60 - 69	60	218	121	56
70 - 79	70	219	155	71
80 - 89	80	238	152	64
90 - 99	90	189	145	77
100	100	227	172	76
	-	1199	802	67

표 9. 10가지 정보를 준 경우의 과신정도

과신정도=0.12

범주(%)	r_t (%)	n_t	V_t	c_t (%)
50 - 59	50	100	52	52
60 - 69	60	194	112	56
70 - 79	70	226	150	66
80 - 89	80	288	195	68
90 - 99	90	181	125	69
100	100	194	139	72
	-	1183	773	65

확신은 증가하여 과신의 정도가 심해질 것이라고 하였는데(Oskamp의 가설), 본 실험에서는 이 가설이 채택되지 못하였다. 실험의 결과를 요약해 놓은 표 8, 9와 그림 6을 보면, 4가지 정보를 받은 경우와 10가지 정보를 받은 경우 과신의 정도는 각각 0.10과 0.12로 별 차이가 없음을 알 수 있다. 더욱이 n_1 의 분포나 v_1 의 분포도 서로 큰 차이가 없다.

정보의 수가 4에서 10으로 증가하였어도 응답의 정확도나 응답에 대한 확신의 정도에 별 변화를 보이지 않았다는 사실은 이들이 확률판정에 어떤 전략을 사용하였는가를 짐작하게 한다. 즉, 많은 양의 정보가 주어진 상황에서 짧은 시

간 안에 직관적인 판정을 내려야 할 때는, 주어진 정보중 중요하다고 생각되는 한 두가지의 정보에 근거하여 판정을 내리고, 나머지 정보는 대부분 무시하거나 약간의 수정을 가하는데 사용한다고 볼 수 있다. 이것은 실제로 실험이 끝난 후에 몇명의 피실험자들과 가진 인터뷰에서도 확인될 수 있었다.

끝으로 정보의 수가 1, 2, 4, 10인 경우를 전부 비교해 보았으나 특기할만한 사항을 발견하지 못하였다.

6. 결 론

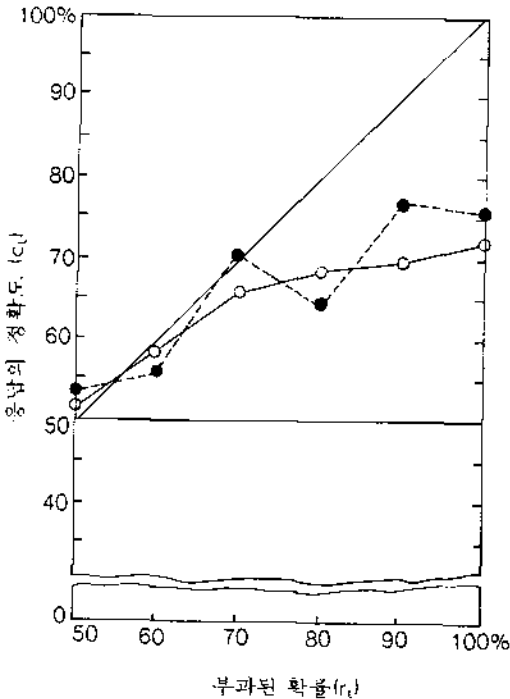
본 연구에서는 직관적인 확률판정에서의 과신 현상이 주어진 정보의 제반특성에 따라 어떤 영향을 받는가를 실험적 방법으로 조사하였다. 실험의 결과는 과신현상이 교정의 측면에서 지배적인 현상이라는 점을 재확인하여 주었다.

또한 정보의 적합성이 낮은 경우와 같이 일반적으로 정보의 가치가 높지 못할 때는 과신현상이 더욱 심해진다는 것을 알 수 있었다.

그러나 정보들간의 중복도는, 최소한 nominal scale로 된 변수를 예측하는 경우, 과신현상에 큰 영향을 미치지 못하며, 정보의 양도 직관적이고 신속한 확률판정에서는 과신의 정도에 큰 영향을 미치지 못한다고 말할 수 있다.

물론 확률판정에서 나타나는 여러가지 행동패턴은 주어진 문제의 유형에 따라 많은 영향을 받는 것이 사실이고, 또 연구의 대상이 인간의 행위 자체이기 때문에 개인차가 심하므로, 본 연구의 결과를 확대하여 일반화 시키는 것은 위험할 것이다.

본 연구에서 얻은 결론도 어디까지나 전반적인 피실험자들의 평균적인 행위에 근거한 것으로서, 다른 사람들과 전혀 다르게 행동한 몇몇 예외적인 피실험자들의 경우를 설명하기는 어렵



- : 4가지 정보를 준 경우
- : 10가지 정보를 준 경우

그림 6. 정보의 수가 4인 경우와 10인 경우의 교정곡선

다. 아마도 이것이 본 연구의 가장 큰 한계점이 아닌가 생각된다.

참고문헌

- [1] Cho, S.K.(1986), *Le Phénomène de Surconfiance dans l'Estimation de Probabilités Subjectives et les Caractéristiques d'Information; These de Doctorat 3-eme Cycle Université d'Aix-Marseille III.*
- [2] Einhorn, H. and Hogarth, R.(1978), "Confidence in Judgement: Persistence of the Illusion of Validity," *Psychological Review*, 85(5), pp.395-416.
- [3] Fischhoff, B., Slovic, P. and Lichtenstein, S.(1977), "Knowing with Certainty: The Appropriateness of Extreme Confidence," *Journal of Experimental Psychology: Human Perception and Performance*, 3(4), pp.552-564.
- [4] Hogarth, R.M.(1980), *Judgement and Choice*, John Wiley and Son, Chichester.
- [5] Jacoby, J., Speller, D.E. and Kohn, C. A.(1974), "Brand Choice Behavior as a Function of Information Load," *Journal of Marketing Research*, 11.
- [6] Kahneman, D. and Tversky, A.(1972), "Subjective Probability: A Judgement of Representativeness," *Cognitive Psychology*, 3, pp.430-454.
- [7] Kahneman, D. and Tversky, A.(1979), "On the Psychology of Prediction," *Psychological Review*, 80(4), pp.237-251.
- [8] Kahneman, D., Slovic, P. and Tversky, A.(1982), *Judgement under Uncertainty; Heuristics and Biases*, Cambridge University Press, Cambridge.
- [9] Keren, G.(1987), "Facing Uncertainty in the Game of Bridge: A Calibration Study," *Organizational Behavior and Human Decision Processes*, 39, pp.98-114.
- [10] Lichtenstein, S. and Fischhoff, B. (1977), "Do those Who Know More Know also More about How Much they Know?" *Organizational Behavior and Human Performance*, 20, pp.159-183.
- [11] Lichtenstein, S. and Fischhoff, B. (1980), "Training for Calibration" *Organizational Behavior and Human Performance*, 26, pp.149-171.
- [12] Lichtenstein, S., Fischhoff, B. and Phillips, L.(1982), "Calibration of Probabilities: The State of the Art to 1980" in Kahneman, D., Slovic, P. and Tversky, A., *Judgement under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge.
- [13] Mehle, T., Gettys, C., Manning, C., Baca, S. and Fisher, S.(1981), "The Availability Explanation of Excessive Probability Assessments," *Acta Psychologica*, 49, pp.127-140.
- [14] Murphy, A.H.(1973), "A New Vector Partition of the Probability Score," *Journal of Applied Meteorology*, 12, pp.595-600.
- [15] Oskamp, S.(1965), "Overconfidence in Case-study Judgements," *Journal of Consulting Psychology*, 29(3), pp.261-265.
- [16] Schum, D., Du Charne, W. and De Pitts, K.(1973), "Research on Human Multi-stage Probabilistic Inference Processes," *Organizational Behavior and Human Per-*

formance, 10, pp.318-348.

[17] Slovic, P., Fischhoff, B. and Lichtenstein, S.(1977), "Behavioral Decision Theory," *Annual Review of Psychology*, 28, pp.1-39.

[18] Spetzler, C.S. and Staël von Holstein, C. -A.S.(1975), "Probability Encoding in Decision Analysis," *Management Science*, 22(3), pp.340-358.

[19] Tversky, A. and Kahneman, D. (1971), "The Belief in the Law of Small Numbers," *Psychological Bulletin*, 76, pp.105

-110.

[20] Tversky, A. and Kahneman, D. (1974), "Judgement under Uncertainty: Heuristics and Biases," *Science*, 185, pp.1073-1078.

[21] Winkler, R.L. and Murphy, A.H. (1968), "Good Probability Assessors," *Journal of Applied Meteorology*, 7, pp.751-758.

[22] Winkler, R.L. and Murphy, A.(1973), "Experiments in the Laboratory and the Real World," *Organizational Behavior and Human Performance*, 10, pp.252-270.