

対訳：リチャード・ニスベット「マインドウェア」からの抜粋：「面接の錯覚：前編」(Richard Nisbett, *Mindware*. Farrar, Straus and Giroux, 2015.)

Even when we're highly knowledgeable about some domain, and highly knowledgeable about statistics, we're likely to forget the concept of variability and the relevance of the law of large numbers. The Department of Psychology at the University of Michigan interviews its top applicants for graduate study before it makes a final decision as to whether to admit a student. My colleagues tend to put substantial weight on their twenty- to thirty-minute interviews with each candidate. "I don't think she's a good bet. She didn't seem to be very engaged with the issues we were discussing." "He looks like a pretty sure thing to me. He told me about his excellent honors thesis and made it clear he really understands how to do research."

The problem here is that judgment about a person based on small samples of behavior are being allowed to weigh significantly against the balance of a much larger amount of evidence, including college grade point average, which summarizes behavior over four years in thirty or more academic courses; Graduate Record Exam (GRE) scores, which are in part a reflection of how much a person has learned over twelve years of schooling and in part a reflection of general intellectual ability; and letters of recommendation, which typically are based on many hours of contact with the student. In fact, college grade point average has been shown to predict performance in graduate school to a significant degree[*1] (a correlation of about .3, which as you'll see in the next chapter is rather modest), and GRE scores also predict performance to about the same degree. And the two are somewhat independent of each other, so using them both improves prediction above the level of each separately. Using letters of

なにかの領域について非常によく知っていて、しかも統計学についてもよく知っていても、分散 [variability] の概念と大数の法則の意義をぼくらはついつい忘れてしまいがちだ。ミシガン大学の心理学部は、大学院での研究を希望する志願者たちの上位者たちに面接して、大学院生として受け入れるかどうかの最終判断を下している。ぼくの同僚たちは、志願者 1 人 1 人とやる 20 分~30 分程度の面接に多大な重みを与える傾向がある。「彼女はあんまり見込みありそうに思わないねえ。さっき議論してた問題にあまり熱意がなさそうだったよ」だとか「私には、彼は実に見所があるように見えるよ。優等生プログラムの最終論文ですぐれたものを書いたって話もしてたし、研究のやり方ってものをよく理解してるのが伝わってきたよ」

ここで問題なのは、行動の小さな標本にもとづいて人物を判断することによって、もっと数量の充実した証拠に比べて不釣り合いに重みづけられるのが許容されているって点だ。そうした証拠には、大学の平均成績や、大学院進学共通試験 (GRE) スコアや、推薦状が含まれる。平均成績には、4 年間に 30 以上の学術コースでみせたふるまいが要約されている。GRE は、12 年間の学校生活でどれだけのことを学んできたのかが反映されている部分もあるし、一般的な知的能力が反映されている部分もある。また、推薦状は典型的にその学生と何時間にもわたる接触に基づいている。実際、大学の平均成績はかなりの程度まで大学院での成績を予測することがわかっている [*1]。また、GRE スコアも、同程度まで成績を予測する。そして、この 2 つはだいたいお互いに独立している。だから、この両方を使えば、個別に使った場合の水準を超えるところまで予測は改善される。推薦状を使えば、予測の精度はさらにもう少し高まる。

But predictions based on the half-hour interview have been shown to correlate less than .10 with performance ratings of undergraduate and graduate students, as well as with performance ratings for army officers, businesspeople, medical students, Peace Corps volunteers, and every other category of people that has ever been examined. That's a pretty pathetic degree of prediction – not much better than a coin toss. It wouldn't be so bad if people gave the interview as much weight as it deserves, which is little more than to let it be a tiebreaker, but people characteristically undermine the accuracy of their predictions by overweighting the value of the interview relative to the value of other, more substantial information.

In fact, people overweight the value of the interview so much that they're likely to get things backward. They think an interview is a better guide to academic performance in college than is high school GPA, and they think an interview is a better guide to quality of performance in the Peace Corps than letters of recommendation based on many hours of observation of the candidate.[*2]

だが、30分程度の面接に基づく予測と学部生や大学院生の成績の相関は 0.10 未満なのが示されている。学生にかぎらず、軍の士官や実業家や医学生や平和部隊志願兵など、これまで調査されたあらゆる範疇の人たちの成績の相関も同様だ。この数字は、予測の度合いとしてはずいぶん悲惨だ——投げたコインの裏表で決めるのと大差ない。人々が面接に相応の重みを与えていればそうひどいことにはならないだろう——つまり、[他の基準で同点になった候補からさらにしぼるための] 延長ゲーム程度にとどめておけばいい。なのに、人間というやつは、他のもっと実質的な情報の価値にくらべて面接の価値を過大評価して、自分の予測精度を低めてしまう。

それどころか、人々は面接の価値を過大評価するあまりに、物事の順序をさかさまにとらえてしまいがちだ。高校の PGA よりも面接の方が大学での学術成績の手がかりとしてすぐれていると考えたり、平和部隊での成績の手がかりとして、数時間にわたる候補者の観察に基づく推薦状よりも面接の方がすぐれていると考えたりする。[*2]