

魔術的相関と仮説形成

脇坂 崇平

理化学研究所脳科学総合研究センター

“Voodoo correlation in social neuroscience”という挑発的なタイトルにてウェブ公開された論文¹が引き金となった一連の議論が、2008年度末より脳科学の一分野にて盛り上がっている。様々な研究グループの反響を引き起こしつつようやく *Perspectives on Psychological Science* に掲載されたこの論文において Vul らは、多くの fMRI 研究に見受けられる不自然なまでに高い（故に“魔術”によるかの様な）相関は、実験データの不適切な取り扱いに起因していると主張した。端的に言えば、仮説形成（“どの”部位が関連しているか）に用いたデータを仮説検証（その部位が”どれ程”関連しているか）に用いる、という一種の循環論法が行われている、というのである。（ところでこの Voodoo Correlation という言葉は、その後クレームがついたのか論文タイトルから除外された。推論と検証に同一データを使うことを NIH の Kriegeskorte, Baker らは”double dipping（ソースの二度漬け）”と呼んでいる。以後こちらをキーワードとして用いる。）実は Vul らの主張そのものが、不適切なデータ（科学者への不完全なアンケート結果）に基づいているのではないかという批判がなされ、本人たちも一部認めたという”オチ”がついている。とはいえ提起された問題自体はいささかも消え去りはしない。Baker らも、主要な科学雑誌に発表された fMRI 研究の少なからずにて、（主張に影響を与える程ではなくとも）double dipping がやはり行われていたと指摘している²。

これは一体いかなる事態なのだろうか？詳細は省くが、double dipping が忍び込む危険性は、被験者数を多くとることが困難で、要素数が膨大、かつ大量のノイズが乗ったデータを扱う fMRI 研究の様な分野では特に大きい。とはいえ、基本的には統計学上の初歩的ミスのはずだ。明確な手続きを踏めば回避できるのではないだろうか。確かに多くの場合はそうである。対処法が存在する。面白いのは、この論争のうちに、現在の科学方法論上排除しようのない類の double dipping が垣間見えているということだ。本発表では一連の論争を出発点とし、その科学方法論上の意味について考察する。

ところで、排除し得ない double dipping とはどのようなものか。例えば以下のやり取りがアナロジーとして成立する。

A と B がジャンケンをして、A が勝つとしよう。ここで、「自分はジャンケンが強いかもしれない。もし二回連続で勝ったなら、ジャンケンに強いといえるだろう。」と A が言う。いささか強引な（実際適切ではない）主張ではあるものの、B は同意することにした。A がおもむろに続ける。「ところで私はすでに一度勝ったので、次も勝ったら私が強いということになる」—B はこれに承諾

しかねたが、A の発言の何が具体的にどう間違っているのかよく分からなかった。さて、B はどのように反論すればよいだろうか？

確率統計の教科書に出てくるような問いである。A は、ある事象（ジャンケンの結果）にあわせる形で蓋然的に自分に都合のよい仮説（「二度連続で勝てば強く、かつ A は強い」）を作り、検証プロセスにその事象自体を取り入れるという誤謬を犯しており、故に「A が強い」という帰結が不当に得られやすくなっている。これもまた、implicit ではあるが **double dipping** の一つと見なすことができる。模範解答は明らかだ。B は”仕切り直し”を要求すればよい。一度目のジャンケンからはじめるというのが妥当である。

だが問題は、A に以下の如く切り返されたときである。「実は最初のジャンケンをする前から、二度勝ったら強く、そして私は強いと考えていた」—この仮説そのものは無時間的なものであるし、こうなると結局、A を信用するか否かという奇妙な話にすり変わってしまうのである。この様な **double dipping** は、完全に隠蔽可能であり、排除しようがない。現象と理論を接続する実験科学に必要な確率統計学上のポリシーというものが、それを A が不誠実にも破っただけだ、という指摘は必ずしも適切ではない。なぜなら、実験データを吟味して秩序を見出し、仮説を形成し、その妥当性を検証する為に最初に吟味したデータも用いるというのは、おそらくはごく普通に行われている行為だからである。そしてその様な煩雑な営為の時系列を整理し、無時間的なものに再配置した後に論文として発表するというのが、我々が現在有する科学のプロトコルである。データを見て仮説を思いついたのか、仮説をもちつつデータを見たのか—これは不問に帰されるが、実のところ帰すべきというよりは、むしろ問いたくても問いやがない、といったほうがいいのではないか。そうして場合によっては、有意でない主張が有意なものとして提示される、といったことが生じ得るのである。逆もまたあるだろう。仮説形成という実際には時間的な営為を無時間的に記述する際に、抜け落ちるもの、新たに加わるものがある、という可能性を我々は忘れてはならない。しかも、仮説を形成する科学者に起因するのか、現象に起因するのかを判断できない形で、である。これは科学的主張の限界というよりは、本質的性質として捉える方が妥当であると私は考える。

科学者の営為とその主張の分離不可能性に関わる上述の議論は、とくに目新しいものではないかもしれない。だが **double dipping** の論争の一部は、この古くから存在する方法論上のテーマに確実に接続した形でなされているし、実際 fMRI 等の解析手法は未だ発展途上なのだ。近年良くも悪くも巷で大量消費され、最新科学の旗頭のような扱いをされがちな「脳科学」の内部でこういった論争が巻き起こっているとは、なかなか興味深いものである。

1. Vul, E., Harris, et. al., Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Psychological Science*. in the press (2008).
2. Kriegeskorte, N., et. al. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 12, 535-40 (2009).