

読心のモデル説と素朴心理学的モデルの制約

藤原 諒祐 (Ryosuke Fujiwara)

東京大学・日本学術振興会

われわれは、他人の考えていることや望んでいることを（ある程度）理解すること、すなわち、心を読むこと（mindreading）ができる。このような能力は、他人の行動の予測・説明のみならず、その解釈や評価、あるいは操作などのために必須となる。つまり、読心能力はわれわれの社会的生活にとって必要不可欠なものである。

われわれにこのような能力が備わっていることから、しばしば、われわれは「心の理論 (Theory of Mind)」をもっているのだ、と表現される (Premack & Woodruff, 1978)。また、時に、われわれが日常的に行う他者行動の予測・説明は「素朴心理学 (folk psychology)」に基づいたものである、と主張される (Churchland, 1981)。読心メカニズムについての「理論説 (theory theory)」は、こうした表現や主張を文字通りに解釈し、われわれの読心実践と科学者の理論的実践とを類比的に捉えようとする。理論説によれば、科学者が科学理論によって自然現象やその背景のメカニズムを理解しようとするように、われわれは素朴理論 (素朴心理学) によって行動やその背景にある心理を理解しようとしているのである。

しかし、理論説が言う「理論」とは具体的にどのようなものなのだろうか。伝統的には、理論を公理系として捉える科学哲学上の見方に従って素朴心理学は解釈されてきた (Maibom, 2003)。例えば、心の哲学の議論の中では、素朴心理学は心的状態と行為にかんする無数の法則的一般化の集合として捉えられてきた (Churchland, 1970; 1981; Lewis, 1972)。類似の見方は、理論説を擁護する心理学者によっても提示されている (Baron-Cohen, 1995)。

一方、科学哲学の文脈では、公理系として理論を捉える見方に代わるものとして、理論をモデルの集まりと捉える見方が提示されている (Giere, 1988; Suppe, 1989 など)。近年、理論説の内部でも、後者の理論観を採用する論者が現れている。こうした論者は、法則的一般化の集合として素朴心理学を捉える見方の問題を指摘し、代替案として素朴心理学をモデルの集まりとする見方を提示する (Maibom, 2003; Godfrey-Smith, 2005; Menzies, 2010 など)。このような「モデル説 (model theory)」によれば、素朴心理学に含まれる知識は、現実の人間行動にかんする法則的一般化についてのものではなく、現実の人間行動のモデルとしての「合理的エージェント」の振る舞いについてのものである (Maibom, 2003)。

モデル説の魅力は、科学理論についての有力な見方を下敷きに加えて、素朴心理学を法則的一般化の集合とする見方の難点を克服しつつ (Maibom, 2003; Menzies, 2010 など)、従来の見方にはない説明上の利点を有していることである。特に、様々なモデルの形式や解釈、使用のあり方が許容されることから、モデル説は他者理解の多様なあり方を捉えうる (Godfrey-Smith, 2005; Spaulding, 2018)。

一方で、モデル説が内実の伴ったプログラムとなるためには、モデルの多様性を許容するだけでなく、様々な場面で用いられるモデルのあり方や特徴を具体的に分析することが必要になる。しかし、素朴心理学において用いられるモデルがどのようなものか、あるいはどのように表象されるのか、という点について詳細な議論がなされているとは言いがたい。

そこで、本発表は、モデル説の妥当性を確認しつつ、素朴心理学的モデルのあり方について考察することを目的とする。そしてそのために、モデル説の論者の1人である P. Menzies の議論 (Menzies, 2010) に焦点を当てる。彼は、素朴心理学をモデルの集まりと捉えることで、素朴心理学が行為の因果的説明と行為の合理的解釈の両方を提供することに説明を与えようとする。本発表の関心にとって重要なことは、彼が、統計学において因果を扱う際の道具立て (構造方程式と因果グラフ) を用いて素朴心理学のモデルを具体的に提示している点だ。発表では、こうした見方が Menzies の議論上の目的を果たすかどうかを検討する。そして、Menzies が提示するようなモデルでは、素朴心理学における心的状態と行為の合理化関係を上手く捉えられないと論じ、そこから妥当な素朴心理学的モデルが充足すべき制約を導く。

主な文献

- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. MIT Press.
- Churchland, P. M. (1970). The Logical character of action-explanations. *The Philosophical Review*, 79(2), 214–236.
- Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *The Journal of Philosophy*, 78, 67–90.
- Giere, R. N. (1988). *Explaining science: A cognitive approach*. The University of Chicago Press.
- Godfrey-Smith, P. (2005). Folk psychology as a model. *Philosopher's Imprint*, 5(6), 1–15.
- Lewis, D. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*, 50(3), 249–258.
- Maibom, H. (2003). The mindreader and the scientist. *Mind and Language*, 18(3), 296–315.
- Menzies, P. (2010). Reasons and causes revisited. In M. De Caro & D. Macarthur (Eds.), *Naturalism and normativity* (pp. 142–170). Columbia University Press.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 4, 515–526.
- Spaulding, S. (2018). *How we understand others: Philosophy and social cognition*. Routledge.
- Suppe, F. (1989). *The semantic conception of theories and scientific realism*. University of Illinois Press.