

研究と社会をつなぐ心理検査・心理診断 — 目的外使用の観点から

早稲田大学文学学術院 教授

小塩真司

「完璧な心理検査」が存在したら

イギリスのコメディ・グループであるモンティ・パイソンは、「空飛ぶモンティ・パイソン」で『殺人ジョーク (The Funniest Joke in the World)』というスケッチ (コント) を演じている。

第二次世界大戦の頃のある日、イギリスの作家が偶然、世界でいちばん面白いジョークを書き上げてしまう。そのジョークはあまりにも面白く、彼は自分が書いたジョークを読み、笑い死にしてしまう。彼の母親もジョークを読んで犠牲になり、事件を捜査した警察官もそのジョークの犠牲になる。このジョークの高い殺傷能力に気づいたイギリス軍は、戦争相手の言語であるドイツ語にジョークを翻訳することを試みる。多くの犠牲を出しつつ翻訳が試みられ、無事に敵国言語に翻訳されたジョークは戦争で絶大な効果を発揮し、戦争は終結する……。

さて、冒頭からモンティ・パイソンの有名なスケッチを長々と紹介した。なぜかというと、心理学で作成されるテストや測定ツールは、時にこの「殺人ジョーク」と同じような社会的現象を生み出す可能性があるからである。

信頼性と妥当性が極限まで高められた、完璧な心理検査が存在することを考えてみよう。検査を受けて出力された結果は、本人が否定しようとも確実な本人像を表現し、個人の行動を制限する。「リーダーに向いていない」という結果が出されれば、本人がいかにリーダーになろうと努力をしても無駄になってしまい、「人に迷惑をかけがち」と出ればそれだけで社会的排除の対象となってしまう。そしてその検査は、人々の管理や統制に利用される。逆説的ではあるが、現実の心理学的

な検査は完全ではないからこそ、われわれは気軽に現実に応用することができるのかもしれない。

なお以下では、ある心理的な特性を客観的に測定する標準化されたツールを検査、心理学的な概念を測定する標準化されていない質問項目群を尺度、個人の心理状態を評価することを診断と呼ぶ。

知能検査と社会的影響

これまでの心理学の歴史の中でも、もっとも社会的に影響力が大きい心理学的検査は、知能検査である。20世紀初頭、アルフレッド・ビネ (Binet, A.) は、シモン (Simon, T.) とともに就学困難な児童を早期に発見し、特別な教育支援を提供するスクリーニングを目的として知能検査を開発した¹。

ビネーらの検査が発表されるとすぐにゴダード (Goddard, H. H.) が知能検査をアメリカ合衆国で紹介し、「精神薄弱 (feeble-mindedness)」のスクリーニング検査として知能検査を用いた。彼は、移民審査局が置かれていたニューヨークに浮かぶエリス島でも知能検査を実施し、多くの移民を精神薄弱者だと断じている²。このような報告は、1924年に制定される移民制限法の科学的根拠にもつながった。これは、心理学的な検査が、国家の政策にも影響を与えたひとつの例だと言える。

1917年、アメリカ合衆国が第一次世界大戦に参戦するにあたり、アメリカ心理学会 (APA) は当時会長であったヤークス (Yerkes, R. M.) を中心として委員会を発足させ、集団式の知能検査であるα検査 (言語式検査) とβ検査 (非言語式検査) を開発した³。この検査によって、大量に徴兵される人々に対して効率的に検

表1 知能検査の目的

目的	例
教育的支援	ビニー式検査による学習困難児の早期発見
社会・政策的利用	ゴダードによる移民審査
軍事・人材選抜	陸軍式の徴兵時検査
臨床・診断利用	ウェクスラー式知能検査
研究利用	標準化検査の代替検査
エンターテインメント	オンラインでの「IQ診断」など

査が実施され、高得点の候補は士官候補に選抜されたり、より責任のある任務への配置が進められたりした。この試みは、現在の人材管理の原型とも言える。

第二次大戦期、病院やリハビリテーション施設で成人患者の知的能力を臨床的に評価する必要が生じていた。ニューヨークのベルビュー病院に勤務していたウェクスラー（Wechsler, D.）は、成人を対象とし、複数の知能領域を測定し知能指数を算出可能な知能検査を開発した。1939年に成人版の検査が開発され、その後、WAIS（成人版）、WISC（児童版）、WPPSI（幼児版）へと展開していくことになる。

標準化された知能検査の代わりに、研究場面では三段論法などの簡易的な課題⁴や認知能力課題、語彙課題などが認知能力の測定のために用いられることが多い。これらは各種の知能検査との間に高い相関関係が報告され、研究目的であれば知能検査の代替検査として無償で利用することができる。

一方で、オンラインで誰もが受験し、自己診断することができる「IQ診断」的なサービスも多数存在する。このような診断は、結果をSNSで拡散し、多くのアクセスを集めることを目的としており、広告収入や別のサービスへの導入、収集したデータを何らかの形で活用しようとする意図が推察される。ここまで知能検査の流れは、表1のように目的を整理することができる。

パーソナリティ検査と社会的影響

次に、パーソナリティに注目して、表2に沿って目的を整理してみたい。まず1919年、Woodworth Personal Data Sheetというパーソナリティ検査が発表された⁵。

現在のPTSDのような症状は「戦争神経症」と呼ばれており、これはBig Fiveモデルの神経症傾向に受け継がれていく。これは第一次世界大戦における兵士たちの神経症リスクのスクリーニングのために開発された、

表2 パーソナリティ検査の目的

目的	例
軍事・スクリーニング	Woodworth Personal Data Sheetによるスクリーニング
臨床・診断利用	MMPIの現場での利用
研究的利用	各種尺度・検査
自己理解・人材開発	MBTIの研修等での利用
社会・産業的利用	採用試験や適性検査
エンターテインメント	オンラインでの性格診断

世界最初のパーソナリティ検査とも言われる。

より広く臨床場面など実践の場で用いられた最初のパーソナリティ検査は、1943年に発表されたMMPI（Minnesota Multiphasic Personality Inventory）である⁶。MMPIは臨床場面や精神医学場面で急速に普及し、世界各国で翻訳され標準化されていった。MMPIは臨床場面で用いられる診断の補助、治療方針の検討、患者やクライエントの特性の変化などを把握するために用いられる検査である。その一方で、論文データベースPsycINFOで「MMPI」が登場する文献を検索すると1万件以上がヒットするように、研究文脈でも盛んに用いられてきた。また、アメリカ合衆国では一部の企業や公的機関が、採用試験や適性の評価にMMPIを利用してきた歴史をもつ。加えて、インターネット上で検索すると、「MMPI風」あるいはMMPIを流用したかのような性格診断を見つけることができる。

MBTI（Myers-Briggs Type Indicator）は、第二次世界大戦期にブリッグス（Briggs, C. K.）とマイヤーズ（Myers, I. B.）によって開発された、ユングのタイプ論を基盤にした自己理解のためのパーソナリティ類型診断ツールである。MBTIは精神疾患の診断や適性の診断でも個人を自動的に診断するものではなく、協会が認めた専門家が主導するセッションで自己理解と他者理解を進める際の補助的な位置づけとなる。そのため、教育現場や企業研修、人材開発の実践現場などで利用してきた。近年日本では「MBTI風の性格診断」が若い世代を中心に流行しているが、これは（もしもMBTIだと主張するなら）本来のMBTIとは異なる利用方法である。

Big Fiveパーソナリティは、20世紀初頭の語彙研究に基礎を置く。1960年代から1980年代にかけて行われた因子分析的研究の蓄積から、多くのパーソナリティ語が「外向性」「協調性（調和性）」「勤勉性」「神経

症傾向」「開放性」という5つの因子に収束することが示され、「Big Five」と呼ばれるようになった。この枠組みは人間のパーソナリティを複数の測定可能な次元の連続体として捉え、個人差を定量化する。1990年代以降、複数の尺度が開発され多くの研究で利用されるとともに、Big Fiveは教育場面や企業の人材管理場面でも活用されてきた。またオンラインでは一般向けのエンターテインメントに近い形でも「性格診断」として流通するようになっている。

どのような検査・調査が存在しているのか

ここまで、いくつかの心理学的な検査や診断の歴史的背景や広がりを見てきた。これらについて、以下のような整理を試みてみたい。

第1に、学術研究用のツールである。心理学の論文の中で、心理尺度や心理検査、測定ツールを開発して論文に掲載することがある。論文に掲載される心理尺度は無料で提供され、研究を目的とする場合には論文を引用することを条件に自由に使用可能である。尺度構成の論文では、質問項目の作成過程、信頼性と妥当性の検討プロセスと結果が記載される。また多くの場合、研究では心理尺度は集団全体の傾向を検討するために用いられ、個人に対して「あなたはこのような心理状態である」といった診断を行うことには適さない。

第2に、専門実践用のツールである。医療や心理臨床、教育の現場で個人の心理的な特徴を理解し、活用するために心理検査が用いられる。こうした心理検査は集団から得られた基準データに基づいて標準化されており、検査対象となる個人に対して検査を実施した際には、集団の中での位置づけから解釈される。検査開発と標準化の手続きにはコストがかかることから、現場で活用される検査の多くは有償で提供される。現場では、公認心理師や臨床心理士などの専門家が検査を実施し、治療や支援に役立てる。また医療現場では各検査に診療報酬点数が定められており、「発達及び知能検査」「人格検査」「認知機能検査その他の心理検査」それぞれについて、難易度に応じて点数が割り当てられている。専門実践用の心理検査は研究に基づくものも多く、成立過程は書籍やマニュアルに記載され、信頼性や妥当性の情報も開示されている。

第3に、組織・管理で利用されるツールである。営利活動の中では、研究の世界とはまた異なった独自の世

界が展開している。企業や団体向けの営利的なサービスとして、各種の適性診断、パーソナリティ診断、研修活動を伴う各種診断が開発され、新規採用や転職、社内の研修、人事に伴う適材適所の判断などに利用されている。また、行政制度においても、選抜や診断、政策決定のために、有償・無償含めさまざまな調査が行われ、政策決定に結びつけられる。これらのツールを心理学の研究論文で見かけることは少ない。この目的で用いられる診断や検査の多くは、営利を生み出す「商品」であることから、検査の作成過程や信頼性・妥当性の情報が開示されるとは限らず、情報を可能な限り開示して透明性を確保しようとする研究活動とは相容れない部分がある。

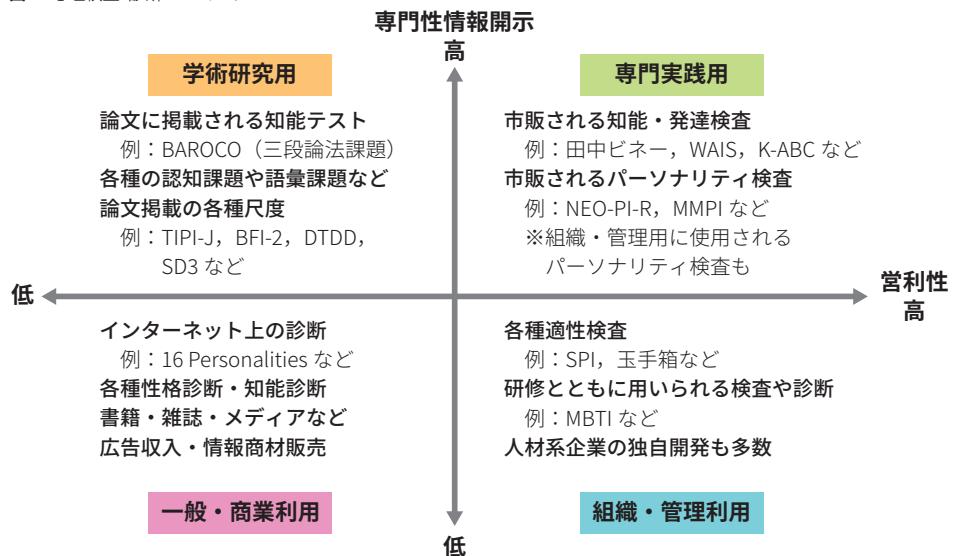
第4に、一般・商業利用のツールである。古くから、書籍やテレビなど各種メディアの中では「心理ゲーム」「心理診断」「心理テスト」といった名称で、エンターテインメントを目的とした診断が提供されてきた。近年ではインターネット上でこの種の診断が提供されており、多くの人々が自己診断に用いている。特に日本では2020年頃から「MBTI診断（実際にはMBTIとは異なるオンライン診断）」が短期間に広く普及した。この種の心理診断は無料で診断結果が出力されるとはいえ、広告収入や情報商材の販売、他の有料サービスへの誘導など営利を目的とすることが多い。また、研究のために開発された心理尺度や検査を無断で掲載する例も散見される。これは、本来の目的から逸脱した心理尺度の利用だと言える。時にこの種のツールは「自己理解を促す」ものとして受け入れられるが、やはり科学的な妥当性よりも娛樂が優先されやすい。

これらの類型は、「営利一非営利」「専門性情報開示高一低」というマトリックスに整理することができる（図1）。もちろん、完全に4つの類型化が成立するわけではない。たとえば図1では記載されていないがYG性格検査や内田クレベリン検査は、専門実践用にも組織・管理用にも用いられている。また、一般・商業利用はマトリックス上の「営利性が低い」場所に位置しているが、実際には多額の利益を上げている診断も存在すると推測される。

心理学は誤用とどう向き合うべきか

パーソナリティ検査や知能検査の歴史には、いわば「誤用」のような現象が各所に見られる。しかし、その

図1 心理検査・診断のマトリックス



内容は多義的である。ここでは「検査や診断の目的」が重要な観点となる。本来、学術研究用に開発された心理尺度は、やはり学術研究に用いるのが望ましい。学術研究用の尺度を組織・管理場面に用いると強い回答バイアスが伴い、結果をゆがめる可能性がある。また専門実践用に用いられる検査は専門家の管理下で結果が解釈されることが望ましく、インターネット上に公開すれば本人の自由な結果の解釈から予想外の弊害をもたらす恐れもある。さらに専門的な実践現場で用いられる検査を興味本位で拡散したり、自己理解のためだからと気軽に自己診断に用いたりすることは、その検査を本当に必要とする場面（たとえば、自分が病気になり検査実施が必要になるような場面）で、診断や治療に役立つ結果が得られなくなる危険性もある。

そして重要なことは、どのような検査であれ「完全なものではない」という点である。これは、健康診断であれ血液検査であれ、心理学者が作成する心理尺度であれ、医療や臨床現場で用いられる検査であれ同じである。検査を本当に必要としている場面では多くの場合、専門家が誤りの可能性も踏まえたうえで結果の解釈や伝達をサポートする。

一方で、インターネット上の性格診断の結果の解釈

には専門家が介在しない。オンラインで診断に回答し「自分はこういう人間だ」と納得しても、その診断内容が誤っている可能性は常に存在しているが、そのことを適切に伝えてくれる専門家はその場にいない。そのような診断を完全なものだと誤解して自分自身の行動を限定してしまうのは、行動の制限や機会の損失につながる可能性すらある。

このように見てくると、心理尺度であれ何らかの診断であれ検査であれ、使用目的を明確にすることが重要である。専門的な論文や書籍には尺度や検査の質問項目が掲載されることも多く、そのまま異なる目的のもとで利用されてしまうリスクを避けることは不可能である。しかし、だからといってそれを放置するのではなく、研究者も応用を試みる人々も検査や診断を受ける側も、それぞれの立場から検査や尺度の使用目的を常に意識し、考え続けることが求められている。



おしお・あつし

2000年、名古屋大学大学院教育学研究科博士課程後期課程修了。博士（教育心理学）。専門はパーソナリティ心理学、発達心理学。中部大学講師、准教授、早稲田大学文学学術院准教授を経て現職。専門はパーソナリティ心理学、発達心理学。著書に『性格診断ブームを問う：心理学からの警鐘』（単著、岩波書店）など多数。

1 Binet, A., & Simon, T. (1916) New methods for the diagnosis of the intellectual level of subnormals. (*L'Année Psych*, 1905, pp.191-244). In A. Binet, T. Simon & E. S. Kite (Trans.), *The development of intelligence in children* (The Binet-Simon Scale) (pp.37-90). Williams & Wilkins Co. **2** Goddard, H. H. (1913) *Am J Ment Defic*, 18, 105-107. **3** Yerkes, R. M. (1948) Psychological examining in the United States Army, 1921. In W. Dennis (Ed.), *Readings in the history of psychology* (pp.528-540). Appleton-Century-Crofts. **4** Shikishima, C., Yamagata, S., Hiraiishi, K., Sugimoto, Y., Murayama, K., & Ando, J. (2011) *Intelligence*, 39(2), 89-99. **5** Garrett, H. E., & Schneek, M. R. (1928) *J Gen Psychol*, 1, 459-471. **6** Hathaway, S. R., & McKinley, J. C. (1943) *The Minnesota multiphasic personality inventory* (Rev. ed., 2nd printing). University of Minnesota Press. ***COI:**本記事に関連して開示すべき利益相反はない。