

「役に立つ」心理学的特性に関する 論点

——心理学の観点から

小塩 真司

(早稲田大学文学学術院教授)

近年、日本のみならず各国で、採用・配置・昇進・育成などの人事意思決定において、適性検査や性格検査を含む心理学的アセスメントの利用が拡大している。同時に、オンラインの無料性格診断も普及し、診断結果の共有が日常化するなかで、娯乐的診断と組織で用いられる検査の境界は利用者にとって曖昧になりつつある。本稿は、心理学的検査が「役に立つ」とは何を意味するのかを、パーソナリティ心理学と測定論の観点から整理することを試みる。まず、心理学的概念は直接観察できない構成概念であり、とくに「コミュニケーション能力」のような実務語は多義的であること、さらに同名概念の同一視 (jingle) や異名概念の別視 (jangle) によって、測定の不一致が生じうることを論じる。次に、妥当性を内容的・基準関連・構成概念の観点から概説し、妥当性を得点解釈と利用の正当性まで含む統合概念として捉える立場についても整理する。さらに、ポジティブ心理学の展開を手がかりとして、概念の望ましさが規範的評価から望ましい結果の予測へと移行してきた経緯を示す。職務成果の予測研究を踏まえ、効果量は実務的含意をもちうる一方で、成果の大部分は他要因に左右され、集団水準の相関は個人の将来を確定しない点を指摘する。最後に、有用性を組織的 (意思決定支援)・個人的 (自己理解)・説明的 (理論的枠組み) の3側面として整理し、妥当性の継続的検証、成果指標の透明化、過度な決定論と自己成就の予言の回避を、責任ある活用条件として提案する。

目次

- I はじめに
- II 心理学的概念と測定
- III 望ましい心理学的概念とは
- IV 予測の有用性とその条件
- V 「役に立つ」心理学的特性とは

I はじめに

近年、日本のみならず各国において、企業の採用活動や人材配置の場面において、心理学的な背景をもつ各種検査の存在感が急速に高まっている (Ployhart, Schmitt and Tippins 2017; Schmidt and Hunter 1998)。いわゆる適性検査や性格検査、ス

トレス耐性やリーダーシップ特性を測定するとされるアセスメントツールは多様化し高度化しており、外部サービスとして提供されるものも少なくない。それらのツールは新卒採用に限らず、中途採用や管理職選抜、さらには組織内の人材育成や配置転換の判断材料としても活用されるようになっていく。人材マネジメントの高度化が求められるなかで、客観的指標に基づく人事にかかわる意思決定への期待が高まっていることは間違いないであろう (Rousseau 2006)。

一方で、こうした組織向けの検査とは別に、オンラインで、無料で提供される性格診断も広く普及している (小塩 2025)。とくに若年層を中心に、自己理解や対人関係の話題として性格タイプ診断

の結果を共有する文化が形成されつつある。SNS上で診断結果を公開することも、珍しいことではない。「自分はこういうタイプである」というラベルが、就職活動やキャリア選択の場面にまで持ち込まれることもある。娯楽的な性格診断と、企業が正式に導入する心理検査とは本来異なるはずであるが、利用者の側から見れば、その境界は必ずしも明確ではない。

こうした状況のもとで、私たちは知らず知らずのうちに、心理学的検査は科学的であり、客観的であり、したがって意思決定に有用である、という前提を受け入れている可能性がある。とくに採用や配置といった重要な判断において、検査結果が補助的に用いられる場合には、その影響は決して小さいとは言えない。しかし、そもそもこれらの検査は何を測定しているのであろうか。また、その測定結果が「役に立つ」とは、どのような意味を指すのだろうか。それは、予測精度が高いことを意味するのであろうか、また意思決定の説明責任を支えることなのであろうか、それとも納得感を提供することなのであろうか。心理学的検査の有用性の意味は、一義的ではないと考えられる。

本稿では、心理学、とくに心理学的個人差を研究対象とするパーソナリティ心理学の立場から、これらの問題を整理することを試みる。心理検査の科学的基盤、測定の限界、そして組織における活用の意義とリスクを検討しながら、「心理学的個人差が役に立つとはどのようなことを意味するのか」という問題について考察する。

II 心理学的概念と測定

心理学的な概念というのは、色や形のように直接知覚できる実体ではなく、理論的に仮定された抽象的な構成概念である。怒りや不安、外向性や勤勉性といった特性は、私たちが日常的に経験する行動や感情の背後にある「傾向」や「性質」として推論されるものであり、それ自体を直接観察することはできない。心理学的概念を扱うということは、なんらかの方法によって観察可能な行動指標や回答反応を通して、背後にあると仮定される構成概念を推定することを指す。

この点は、実務場面においても同様である。たとえば、人事採用の際に多くの企業が求める資質に「コミュニケーション能力」がある（株式会社帝国データバンク 2022）。しかしながら、「コミュニケーション能力」とは何を指すのか、その定義や構成要素が明確に共有されているとは限らない。対人場面での発話の明瞭さを意味するのか、相手の感情を読み取る共感性を含むのか、あるいはチーム内での調整能力や説得力、他者の意見を傾聴する姿勢までを含むのかによって、その内容は大きく異なる。ところが実際には、「コミュニケーション能力が高い人材が望ましい」といった表現が、定義の明確化を経ないまま用いられることが少なくない。その結果、同じ語を用いていても評価者ごとに想定している内容が異なり、評価基準が不明瞭なままに運用される可能性がある。

藤本・大坊（2007）が整理しているように、コミュニケーション・スキル（学習や教育、トレーニングによって変容することを前提とした概念を、能力よりもスキルと呼ぶことが多い）は単一の内容を指すものではなく、対人的な認知、自己表現、関係維持行動など複数の下位側面から構成される、多面的な概念である。「コミュニケーション能力」という語は、日常的にはひとつの資質であるかのように扱われるが、その内容は理論的立場や測定方法、また文脈によって大きく異なるものだと言える。このことは、企業が求める「コミュニケーション能力」が、具体的に何を意味するのかを明確にしないままでは、評価基準の曖昧さや測定の不一致が生じる可能性を示唆している。

このような概念の多義性に関連して、心理学の測定論において古くから指摘されてきた問題に、いわゆる“jingle fallacy”と“jangle fallacy”がある（Kelley 1927；Thorndike 1904）。前者は、同じ名称が用いられているという理由だけで、それらが同一の構成概念を測定していると誤認してしまう誤謬である。そして後者は、異なる名称が与えられているという理由だけで、実際には類似した構成概念であるにもかかわらず別のものであるとみなしてしまう誤謬を指す。

たとえば、複数の企業がそれぞれ独自の「コミュニケーション能力尺度」を用いていたとして

も、それらが測定している内容が同一であるとは限らない。逆に、「対人スキル」「協働性」「社会的コンピテンス」などといった、異なる名称が与えられていても、実質的には大部分が重複した特性を測定している可能性もある。このような概念と名称のずれは、評価基準の整合性を損ない、測定結果の比較可能性を低下させる要因となる。したがって、企業実務において心理学的概念を用いる際には、その概念の名称がどれだけ普及しているのか、また印象的で好ましくそれらしい名称であるのかという観点だけではなく、その背後にある構成概念の定義と測定内容の妥当性を慎重に検討する必要がある。

心理学的概念について測定を試みる上では、妥当性の検討が不可欠である。妥当性とは、ある検査（テスト、測定尺度）が、本当に測定しようとしている構成概念を適切に測定していると言えるのかという問い、さらには結果をどのように解釈し利用するのかを検討することである（American Educational Research Association, American Psychological Association and National Council on Measurement in Education 2014）。この問題は、古くから心理学的な測定における中心的なテーマであり、Cronbach and Meehl (1955) は、観察不能な構成概念を測定する際には、理論的な概念間のネットワークの中で当該概念の位置づけを明確にし、関連する変数との間の体系的な関係によって妥当性を検証する必要があることを指摘した。その後、妥当性の考え方は、内容的妥当性、基準関連妥当性、構成概念妥当性といった複数の側面に整理されて理解されるようになった（小塩 2025）。

内容的妥当性とは、測定している内容が、評価されるべき概念の範囲の情報を過不足なく含んでいることを指す（三枝 2024）。たとえば、「コミュニケーション能力」を測定する際に、他者に自分の意見を伝える自己主張性のみを測定するのであれば、「コミュニケーション能力」の広範な内容の一部しか含んでいないことになる。また、「コミュニケーション能力」を測定する際に「コミュニケーションとは何か」という知識について尋ねるのであれば、実際にうまくコミュニケーションを行うことができるかどうかとは異なる側面を測

定することになる。これらの例は、いずれも内容の妥当性の観点から問題を含むことになる。

また基準関連妥当性とは、測定された内容となんらかの基準、あるいは予測の対象となる外部指標との関連の程度を検討することを指す（三枝 2024）。心理学的な測定とほぼ同時期に得られる外部の基準との関連を検討することを併存的妥当性、心理学的な測定よりも将来の基準との関連を検討することを予測的妥当性という。たとえば、新たな「コミュニケーション能力」を測定する検査を実施するとともに、上司や同僚による多面的評価（360度評価）で得られた得点との関連を検討し、自己評価による「コミュニケーション能力」の高さが多面的評価における周囲からの評判と関連するかどうかを検討することは、併存的妥当性の検討に相当する。また、入社時に測定された「コミュニケーション能力」が、入社後の営業成績や顧客満足度、リーダーシップの発揮や周囲からの信頼度などを予測することができるかどうかを検討することは、予測的妥当性の検討に相当する。

構成概念妥当性とは、まさに理論的に仮定された概念が、測定された内容にどの程度反映するかを指す。とくに研究の中では、ある概念を測定する尺度の得点が、他の概念を測定する得点とどのような関係を示すか、理論的な仮定に沿った結果を示すかどうかを検討される。構成概念妥当性の検討方法として重要なものに、収束的妥当性と弁別的妥当性がある。収束的妥当性は、理論的にある概念と別の概念間に密接な関係が想定される際に、それぞれの概念を測定した場合でも相関関係が観察されることを指す。一方で弁別的妥当性は、概念間の関係性があまり認められない、あるいはほぼ無関係であると想定される場合に、それぞれの概念を測定した指標間でも低い関連または無関連が観察されるかを検討することを指す。たとえば「コミュニケーション能力」という概念内容からは、Big Five パーソナリティのうち、活動性や社交性を意味する外向性や、他者に対するやさしさや他者を優先する志向性を意味する協調性との関連が予想される。実際に測定された得点間の関連の検討においても、これらの変数間にあ

る程度の高さの正の相関関係が認められることが、収束的妥当性の検討にあたる。また、概念的な内容から、「コミュニケーション能力」は、論理的な思考力や一般的な認知能力とは高い関連を示すとは考えにくい。したがって、実際に「コミュニケーション能力」の得点と知能検査の結果や認知能力検査の得点との間の関連が、それほど高くない相関係数を示す、あるいはほぼ無関連になるのであれば、弁別的妥当性が認められることになる。

近年、妥当性を複数の種類に分けるという理解は、妥当性の理解として必ずしも十分ではないという議論もある。Messick (1995) は、妥当性を全体的な単一の統合的概念として捉え、測定結果の解釈とその利用の正当性まで含めて妥当性について検討すべきであるとした。妥当性とは、単に統計的な相関関係がどのようになるのかを問題にするだけでなく、得点の解釈が理論的また実証的な根拠に十分に支えられているのか、さらにその測定の利用が社会的また倫理的に正しいものと言えるか、ということも含めた、包括的な概念なのである。

測定の妥当性をめぐっては、さらに再定義も試みられている (Borsboom, Mellenbergh and van Heerden 2004)。ある検査やテストが妥当であるためには、想定される構成概念が実在しており、かつその概念の変化が測定された指標の変化を因果的に生じさせる必要がある。この観点からすると、「コミュニケーション能力の高さ」から「コミュニケーション能力の検査得点の高さ」へ、という明確な因果の経路が証明されない限り、この「コミュニケーション能力検査」の妥当性について判断することができないことになる。この考え方は、単に測定された内容と他の変数との関連を検討するだけでなく、より厳しい検討を要する理論的枠組みを提示する。

心理学的な測定における妥当性とは、ある心理学的な構成概念が理論的にどのように定義され、その定義にもとづいて作成された測定指標が、想定されたとおりの意味をもつ得点を生み出していると言えるのかを、多面的な証拠から検証していく営みである。妥当性とは、測定された内容が

「当たるかどうか」という問題ではなく、得られた得点をどのように解釈するか、その解釈をどこまで一般化することができるのか、得られた得点の利用方法は適切なのか、という一連の問いを包括的に検討することを指す。

Ⅲ 望ましい心理学的概念とは

パーソナリティ心理学の歴史のなかで、英単語“character”と“personality”の区別は、望ましい心理学的個人差について考える際のひとつのポイントとなる。Allport (1937) は書籍の中で、“character is personality evaluated, personality is character devaluated”と述べている。これは、「character は評価された personality であり、personality は脱評価された (評価を取り除いた) character である」ということを意味する。すなわち、character という単語が表現するのは、道徳的・倫理的・規範的な評価を加えた人間の個人差であり、personality という単語が表現するのは、そのような評価を取り除いた人間の個人差であることを指すのである。日本語においては、「人格」と「性格」という対比がこれらに相当する。たとえば、「あの人は人格者だ」という表現をすることはできるが、「性格」に「者」をつけて「性格者だ」と表現することはない。このことは、「人格」という単語には望ましい、道徳的で規範的な意味が含まれているのに対し、「性格」という単語には含まれていないことを意味する (小塩 2010)。この点を考慮すると、“character”は「人格」、「personality」は「性格」に近い意味内容をもつと考えられるが、実際には「人格」と「性格」という言葉は、それぞれ言い回しの中で柔軟に使用されている。

心理学的な個人差を研究する心理学の分野が「パーソナリティ心理学 (personality psychology)」と呼ばれる背景には、この“character”と“personality”の区別の経緯もある。すなわち、「パーソナリティ心理学」という名称には、価値中立的な心理学的個人差を研究対象とする、という考え方が含まれているのである。

当然ながら、パーソナリティ心理学の歴史のな

かでは、望ましい・望ましくない心理学的個人差を研究対象としてきた歴史がある。たとえば Allport 自身、宗教性や価値観、成熟した人格といったテーマを研究で扱ってきた (Allport 1961)。また、臨床心理学や精神医学とパーソナリティ心理学との接点では、神経症傾向や反社会的傾向など、個人または社会にとって望ましくない心理学的個人差傾向が研究対象となってきた。とくに、抑うつや不安など精神病理学的に重要な概念や、1980 年代後半から注目されたパーソナリティ症 (パーソナリティ障害) に関連する研究は、多くの研究者が取り組んできた研究テーマの一例であろう。パーソナリティ心理学の研究領域では、価値中立であることを前提としながらも、特定の研究対象や概念的定義の段階で、一定の範囲で価値判断が入り込む余地はあったと考えられる。

この流れに大きな変化をもたらしたのは、1990 年代に生じたポジティブ心理学のムーブメントであった。1998 年にアメリカ心理学会の会長に就任した M. E. P. Seligman は、それまで心理学が主に精神疾患や機能不全の理解と治療を重視してきたことに対して批判的な立場を取り、人間の強みや徳、幸福などポジティブな側面を体系的に研究することの必要性を訴えた (Seligman and Csikszentmihalyi 2000)。

この研究の流れの中で、精神疾患の分類マニュアルに対応するような、ポジティブな精神機能の整理が試みられた。そこでは人間の強み (長所; strength) が注目され、強みを測定する測定尺度として Values In Action Inventory of Strengths (VIA-IS) も開発された (Peterson and Seligman 2004)。この開発に際しては、古典、哲学的議論、宗教などを包括的に検討し、人類に普遍的と考えられる 6 つの美德 (virtue) と、それらを具体化する 24 の強みを整理した (大竹ほか 2005)。ここで整理された 6 つの徳は、知恵と知識、勇気、人間性、正義、節度、超越性である。この枠組みは、“character strength”とも呼ばれ、「強み」「性格的強み」などと訳される。ここで“character”という単語が用いられている点は、象徴的である。価値中立であることに重きをおくことで“personality”という単語を用いてきた

パーソナリティ心理学に対して、「望ましい概念」を明示的に扱う研究の枠組みが提示されたことを意味する。

強みとして整理されたもののうちとくに 6 つの美德は、これまでの人類史においてそれ自体が「良いもの」として論じられてきたものが中心となっている。しかしながら一方で、現代の心理学は、実験や調査など科学的な手法を用いて研究が進められる。

科学的な営みの中で研究が行われると、次のような手続きが行われるようになっていく。それは、「どうしてその概念は良いものと言えるのか」という疑問に対する回答を試みようとする動きである。たとえば 6 つの美德のうち「勇気」については、測定尺度が開発され、日本語化もされている (下司・吉野・小塩 2023)。「勇気」という概念に対して、「勇気尺度」という用具で測定が可能になると、勇気尺度で測定された内容が「どのように良いと言えるのか」を検討していくことが可能になる。勇気得点が高い人物は低い人物に対して学業成績が良いのか、良好な人間関係を営む傾向にあるのか、職業パフォーマンスが良好であるのか、収入が多いのか、そして心身ともにより健康であるのかといった研究である。このような試みは、他の「望ましい」とされる心理学的な概念についても進めていくことが可能になる。

心理学的概念を理論的に言葉で定義することを概念的定義という。一方で、測定方法によって概念が定義されることを、操作的定義という (Cronbach and Meehl 1955)。研究の中で心理学的概念を測定する方法が整備され、概念が操作的に定義されるようになると、「その概念がどのような結果を予測するのか」が検討されるようになる。そこでは、ある心理学的概念が「望ましい」「良い」とされるかどうかは、その測定された概念が何を予測するのか、何に関連するのかという、経験的な証拠によって判断されるようになる。

これは、科学的な研究を進めることで、概念の「望ましき」の基準が、倫理的・規範的な水準から、予測的妥当性という経験的な水準へと移行することを意味する。ある心理特性が「良い」とされるのは、それが概念的に望ましと論じられるか

らではなく、「より良い結果をもたらすことが多い」からなのである。このとき、概念の価値は、その概念がどれくらい将来の適応や成果を説明し、予測するかという、統計的な結果に依存することになる。

さて、ここで「より良い結果」を予測する心理学的概念が「良いもの」だとしても、何をもって「より良い結果」だとみなすことができるのだろうか。そこには、結果を設定する段階で、すでに価値判断が行われているのである。「より良い結果」には、学業成績や学歴、職業的成功や収入、主観的幸福感や生活満足度、心身の健康や生死など多岐にわたるものが設定される。しかし、どのような結果を設定するかによって、その内容は微妙に異なる。この中には、個人にとって望ましい結果もあれば、社会全体にとって望ましいが個人にとっては必ずしも望ましいとは言いきれないものも含まれる。心理学的概念の良さが予測可能性によって決定されるとはいえ、その予測される内容の選択自体が、何に価値を置くかによって左右されるのである。

それでもなお、予測という考え方自体は、心理学的概念の価値を検討する上で、重要な基準を提供すると言える。予測可能性は、再現性と数量的評価を伴い、心理学的概念を科学的に検討する上で非常に重要な方法と言えるからである。ある心理学的特性が多く文化や年齢層、社会的階層において一貫して適応的な指標を予測するのであれば、その概念は理論的にも意義をもち、現実社会における応用可能性についても大きく期待されることになる。

以上のことから、パーソナリティ心理学において「望ましい心理学的概念」とは、道徳的に望ましいとされる概念であるというよりも、「望ましい結果を予測しうる概念」という意味合いが大きいものであらうと考えられる。この観点は、さまざまな概念を応用的に利用する場合にも重要なものである。

IV 予測の有用性とその条件

ここまで、心理学的概念の望ましが、道徳

的・規範的な価値判断によってというよりも、主として「望ましい結果を予測しうるかどうか」という経験的な基準によって判断されるようになってきたという経緯を説明した。人材管理の領域においても、心理学概念の予測可能性が重要な観点になると考えられる。

人材の採用や配置、昇進、育成といった人事に伴う意思決定というものは、いずれも将来を予測する判断を意味する。しかも、限られた情報のもとで、不確実な将来を予測することを強いられる状況にある。このような状況下で、個人の認知的な能力や各種のスキル、そしてパーソナリティ特性が、職業上の成果や実績、組織への適応を一定程度予測することを前提として、人材管理に利用されている。これは、人材管理に用いている心理学的概念や心理的特性が、将来を予測するという前提に立つ行為だと言えるのである。

しかし、ある心理学的概念が将来を予測することができる、という前提に立つということは、用いた心理学的な検査の結果によって、個人の将来が左右されるということも意味する。では、この将来予測というのは、どれくらいの精度で行われるものなのであらうか。

これまで心理学の研究の中では、職務遂行や職業上の達成などを心理学的変数で予測する研究が、蓄積されてきている。たとえば、Schmidt and Hunter (1998) は、過去 85 年間にわたる研究をメタ分析で統合することで、一般的認知能力が職務遂行結果を予測することを報告した。ここで報告された統合後の相関係数は、 $r = .50$ 前後であり、比較的高い水準の予測力だと言える。また、勤勉性をはじめ各種のパーソナリティ特性については、低い水準から中程度の予測力 ($r = .20$ から $.30$ 程度) を示すことが報告されている。この論文では、知的な課題とそれ以外の心理的特性（いわば非認知能力）を組み合わせることで、予測力が向上する可能性が示されている。

また、Sackett et al. (2022) は、人事選抜研究における妥当性係数の推定方法を再検討し、とくに従来のメタ分析において用いられてきた範囲制限 (restriction of range) の補正が体系的に過大であった可能性を指摘した。従来の研究では、実際

の採用場面においてすでに選抜が行われているために観察される相関が過小評価されているとみなし、統計的補正を加えることが一般的であった。しかし、Sackettらは、母集団全体の分布を仮定して機械的に補正を行うことが、現実の選抜過程を十分に反映していない場合があることを示し、補正後の妥当性係数が従来報告されてきた値よりもやや低く見積もられる可能性を報告している。この再分析は、心理学的変数が職務成果を予測するという基本的結論自体を否定するものではないが、予測力の水準が推定方法に依存すること、そして効果量の解釈には慎重さが求められることを示唆している。ただし、Sackettらによる再分析を経ても、人事選考に有効な結果に対して、一般的認知能力は $r=.30$ 程度、勤勉性は $r=.20$ 程度の予測力をもつことが示されている。

さて、これらの研究で示されている効果量（相関係数の値）は、どの程度だと言っているのでしょうか。相関係数 $r=.30$ は、一見すると決して大きな値には見えないかもしれない。しかし、分散説明率に換算すれば約9%に相当する。すなわち、職務遂行の個人差のうち約1割が、当該の心理学的変数によって統計的に説明されることを意味する。単一の心理変数で将来の職務成果の1割前後を説明できるということは、社会科学の領域においては決して小さな効果とは言えない。この効果量の大きさは、医学における救命処置の効果に匹敵するほどの大きなインパクトを持つのである（Meyer et al. 2001）。とくに、多数の応募者の中から相対的に将来の成功確率が高い人物を選抜するという目的においては、この程度の関連の大きさであっても、組織全体として見れば無視できない累積的影響をもつ可能性がある。

しかし同時にこの数値は、職務の成果の大部分が、当該の心理変数以外の要因によって大きく左右されることも意味する。職務の成果は、それまでのどのような経験を積んできたか、どのような組織の中で働くか、上司との関係や同僚との関係、さまざまな巡り合わせや偶然など、多種多様な要因から影響を受ける。たとえば、高い認知能力をもつ人であっても、明確な目標が与えられない職場や、心理的安全性の低い環境では十分に力を発

揮できない可能性がある。逆に、測定された心理特性の得点が平均的であったとしても、適切な訓練や支援、良好なチームワークに恵まれることで、高い成果を上げる場合もある。すなわち、心理学的変数は職務遂行の「一部」を説明するにすぎず、環境との相互作用の中でその効果が発現するとも言えるのである（Tett and Burnett 2003）。ここには、個人特性と状況要因の交互作用という、古くから議論されてきた基本的な問題がある。

さらに、相関関係は、集団における統計的な関連を示す指標に過ぎない。これは、ある個人の将来を確実に予測することを意味するわけではない。ある心理学的な得点と職業上の成功の程度との間の相関係数が $r=.30$ であるときの散布図を描いてみると良い。散布図に描かれる点は右上がりの傾向を示すであろうが、同時に広く散らばっているはずである。高得点者の中にも低い成果にとどまる者が存在し、低得点者の中にも高い成果を達成する者が存在する。統計的予測はあくまで確率の差を示すものであって、個人単位での運命を確定するものではない。そうであるにもかかわらず、選抜や昇進の場面では、その数値があたかも個人の将来に重大な影響を及ぼすかのように用いられる。ここには、集団水準の知見を個人水準の意思決定に適用するという、推論上の飛躍が存在する。この点を十分に理解しないまま数値のみを絶対視するのであれば、過度に予測可能性を見積もることにつながる。

さらには、何をもって職務遂行上「望ましい結果」だとみなすのかという問題もある。職務遂行の基準は、売上高や生産性といった定量的指標だけではない。チームへの貢献、組織文化への適応、倫理的な行動、長期的な人材育成への寄与など、多面的な側面が存在する。ある評価基準を採用すること自体が、特定の価値観を前提としているのである。たとえば短期的な成果を重視する組織と、長期的な成長や協働を重視する組織とでは、「望ましい人材像」は異なるであろう。したがって、心理学的概念の予測力を論じる際には、どの成果指標を用いているのかを明確にする必要がある。

V 「役に立つ」心理学的特性とは

本稿ではこれまで、心理学的特性の有用性を主として予測可能性の観点から検討してきた。しかし、「役に立つ」という語が意味する内容は、単に将来の成果をどの程度予測できるかということの意味するだけではない。心理学的特性は、組織にとって役に立つ場合もあれば、個人にとって役に立つ場合もあり、さらには現象を理解し説明する枠組みとして役に立つ場合もある。したがって、「役に立つ」心理学的特性を論じるためには、有用性について多元的に整理する必要がある。

第1の有用性は、「組織的有用性」とでも呼べるものである。これは、採用、配置、昇進、育成といった組織の意思決定において、心理学的特性がどの程度、実務的価値をもつかという観点である。組織にとっての有用性は、主として予測的妥当性と関連する。すなわち、当該の特性が職務遂行、離職率、チーム適応、安全行動などの指標を一貫して予測しうるのであれば、それは組織的観点から有用であると言える (Sackett et al. 2022 ; Schmidt and Hunter 1998)。

しかし、組織的有用性は単に相関係数の大きさだけで決まるものではない。まず、コストと便益のバランスがある。測定にかかる時間や費用、運用の負担に対して、どれほど意思決定の質が改善されるのかが問題となる。また、公平性と説明責任の問題がある。特定の集団に不利な影響を及ぼさないか、意思決定の根拠として社会的に正当化できるかという観点が不可欠である。さらに、状況との適合性がある。ある特性が有効であるかどうかは、職務内容や組織文化によって変化する。したがって、組織的有用性とは、特定の文脈において意思決定を改善するか、という文脈に依存する概念であるとも言える (Tett and Burnett 2003)。

第2の有用性は、「個人的有用性」である。これは、心理学的特性の測定が、個人の自己理解、意思決定、成長にどのような利益をもたらすのかという観点である。オンラインの性格診断の普及が示しているように (小塩 2025)、人々は自らの特性に関する情報を通じて、自分がどのような傾

向をもつのかを知りたいと考えている。このとき、特性の測定は、単なる予測装置ではなく、自己認識の枠組みとして機能する。

個人的有用性は、必ずしも高い予測的妥当性を前提としない。たとえば、自らの勤勉性や外向性の傾向を知ることが、日常的な行動選択や人間関係の調整に役立つ場合がある。あるいは、自分の弱みを認識することが、補完的な戦略をとる契機となる場合もある。この意味で、心理学的特性は「未来を当てる」ためというよりも、「現在を理解し、行動を調整する」ための情報として役に立つ。

もっとも、個人的有用性にも注意が必要である。特性ラベルが固定的な自己理解を生み、「自分はこういうタイプだから仕方がない」という決定論的思考につながる危険もある。したがって、個人的有用性が実現するためには、特性を不変の本質としてではなく、「傾向」や「確率的な偏り」として理解することが前提となる。

第3に、「説明的有用性」である。これは、心理学的特性が、人間行動や社会現象を体系的に理解するための概念枠組みとしてどの程度有効であるかという観点である。特性概念は、個々の行動をばらばらに捉えるのではなく、一定の範囲の一貫性をもつ傾向として整理することを可能にする。たとえば、ある人物の協調性が高いという特性に関する記述は、対人場面における行動の安定したパターンを説明する理論的枠組みを提供する。

説明的有用性は、予測的妥当性と関連しつつも、相互に区別される。ある特性が必ずしも高い予測力を示さない場合であっても、理論的整合性があり、関連変数との体系的関係をもつのであれば、行動理解の枠組みとして価値をもつ。Cronbach and Meehl (1955) が指摘したように、構成概念は理論的ネットワークの中で位置づけられることによって意味をもつ。説明的有用性とは、この理論的ネットワークを通じて、人間行動の背後にある構造を明らかにする可能性がある。

さらに、説明的有用性は、研究の累積性とも関係する。概念が明確に定義され、測定が標準化されることで、異なる研究結果を比較し統合することが可能になる。この意味で、心理学的特性は、個別の意思決定を超えて、知識体系の構築に寄与

する。

以上のように、「役に立つ」心理学的特性とは、単に予測力の大小によって評価されるものではない。組織にとっての意思決定支援という有用性、個人にとっての自己理解という有用性、そして理論的説明枠組みとしての有用性が重なり合いながら存在している。これらはいずれも重要であるが、同時に相互に矛盾する関係が生じる場合もある。ある特性が組織的には有用であっても、個人にとっては弊害をもたらすケースがあるかもしれない。あるいは、理論的には精緻であっても、実務的には扱いにくい概念である場合もある。

このように有用性が多元的であることから、心理学的特性を社会的実践の中で用いる際には、いくつかの基本的条件が満たされている必要があると考えられる。まず求められるのは、妥当性の確保である。ある特性が何を意味し、どのような理論的前提のもとで定義され、いかなる測定手続きによって得点化されているのかが明確でなければならない。さらに、その得点が想定された構成概念を適切に反映していること、関連変数との関係が理論的予測と整合していること、そしてその利用が社会的に正当化可能であることを、多面的な証拠から検証を重ねる必要がある。妥当性は一度確認すれば十分であり、それ以上検討する必要が無い、という性質のものではない。それは、その測定が利用される文脈の変化に応じて、継続的に点検されるべき性質のものである。

また、心理的な変数が予測する成果指標の十分な検討と透明性が求められる。心理学的特性の「予測力」を論じる際には、何をもって望ましい結果とみなしているのかが明確にされなければならない。売上高、生産性、離職率、チーム満足度、倫理的行動など、成果指標は多様であり、どの指標を採用するかは価値判断を含む選択である。この選択が曖昧なままに置かれるとすれば、心理特性の有用性の評価もまた、曖昧で暗黙の前提に立つ価値観に依存することになる。したがって、どの成果を重視し、なぜそれを重視するのかを明示することは、心理学的特性の利用を正当化するうえで不可欠である。

そして、過度な決定論の回避も重要な観点であ

る。心理学的特性は、確率的な傾向を示す指標であり、個人の将来を確定するものではない。そうであるにもかかわらず、得られた得点が、何らかの固定的な能力や本質を示すかのように解釈されるならば、自己成就的予言 (Jussim and Harber 2005) や機会の制限 (小塩 2025) を生み出す危険性がある。特性は状況との相互作用の中で何らかの結果をもたらすものであり、環境の設計や支援の在り方によってその表れ方は変化しうるものでもある。この前提について、常に認識し続けることが、心理学的な特性を社会の中で適切に用いるためには必要である。

これらの観点は、多元的な有用性をもつ心理学的特性を、社会の中で責任ある形で活用するための基本的な観点だと言える。心理学的特性が「役に立つ」ためには、その有用性の側面を見極めるだけでなく、その利用を支える理論的・実証的・倫理的な背景について、常に点検し続ける姿勢が求められる。

参考文献

- 大竹恵子・島井哲志・池見陽・宇津木成介・クリストファー・ピーターソン・マーティン、E. P. セリグマン (2005) 「日本版 生き方の原則調査票 (VIA-IS: Values In Action Inventory of Strength) 作成の試み」『心理学研究』76 巻 5 号, pp. 461-467.
- 小塩真司 (2010) 『はじめて学ぶパーソナリティ心理学——個性をめぐる冒険』ミネルヴァ書房.
- (2025) 『性格診断ブームを問う——心理学からの警鐘』岩波書店.
- 小塩真司編著 (2024) 『心理尺度構成の方法——基礎から実践まで』誠信書房.
- 株式会社帝国データバンク (2022) 「企業が求める人材像アンケート——企業が求める人材像, 「コミュニケーション能力が高い」がトップ」. Retrieved from <https://www.tdb.co.jp/report/economic/ayashiflfts6/> (2026 年 2 月 22 日アクセス)
- 下司忠大・吉野伸哉・小塩真司 (2023) 「日本語版勇気尺度の作成——項目反応理論を用いた尺度構成」『心理学研究』94 巻 1 号, pp. 43-53.
- 藤本学・大坊郁夫 (2007) 「コミュニケーション・スキルに関する諸因子の階層構造への統合の試み」『パーソナリティ研究』15 巻 3 号, pp. 347-361.
- 三枝高大 (2024) 「心理尺度の妥当性」小塩真司編著『心理尺度構成の方法——基礎から実践まで』誠信書房, pp. 89-100.
- Allport, G. W. (1937) *Personality: A Psychological Interpretation*, New York: Henry Holt.
- (1961) *Pattern and Growth in Personality*, Holt, Reinhart & Winston.
- American Educational Research Association, American Psychological Association and National Council on Measurement in Education (2014) *Standards for Educational and Psychological Testing*, American Educational Research

- Association.
- Borsboom, D., Mellenbergh, G. J. and van Heerden, J. (2004) "The Concept of Validity," *Psychological Review*, Vol. 111, No. 4, pp. 1061-1071.
- Cronbach, L. J. and Meehl, P. E. (1955) "Construct Validity in Psychological Tests," *Psychological Bulletin*, Vol. 52, No. 4, pp. 281-302.
- Jussim, L. and Harber, K. D. (2005) "Teacher Expectations and Self-Fulfilling Prophecies: Knowns and Unknowns, Resolved and Unresolved Controversies," *Personality and Social Psychology Review*, Vol. 9, No. 2, pp. 131-155.
- Kelley, T. L. (1927) *Interpretation of Educational Measurements*, World Book Co.
- Messick, S. (1995) "Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning," *American Psychologist*, Vol. 50, No. 9, pp. 741-749.
- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszy, T. W. and Reed, G. M. (2001) "Psychological Testing and Psychological Assessment: A Review of Evidence and Issues," *American Psychologist*, Vol. 56, No. 2, pp. 128-165.
- Peterson, C. and Seligman, M. E. P. (2004) *Character Strengths and Virtues: A Handbook and Classification*, Oxford University Press; American Psychological Association.
- Ployhart, R. E., Schmitt, N. and Tippins, N. T. (2017) "Solving the *Supreme Problem*: 100 Years of Selection and Recruitment at the *Journal of Applied Psychology*," *Journal of Applied Psychology*, Vol. 102, No. 3, pp. 291-304.
- Rousseau, D. M. (2006) "Is There Such a Thing as 'Evidence-Based Management'?" *Academy of Management Review*, Vol. 31, No. 2, pp. 256-269.
- Sackett, P. R., Zhang, C., Berry, C. M. and Lievens, F. (2022) "Revisiting Meta-analytic Estimates of Validity in Personnel Selection: Addressing Systematic Overcorrection for Restriction of Range," *Journal of Applied Psychology*, Vol. 107, No. 11, pp. 2040-2068.
- Schmidt, F. L. and Hunter, J. E. (1998) "The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings," *Psychological Bulletin*, Vol. 124, No. 2, pp. 262-274.
- Seligman, M. E. P. and Csikszentmihalyi, M. (2000) "Positive Psychology: An Introduction," *American Psychologist*, Vol. 55, pp. 5-14.
- Tett, R. P. and Burnett, D. D. (2003) "A Personality Trait-based Interactionist Model of Job Performance," *Journal of Applied Psychology*, Vol. 88, No. 3, pp. 500-517.
- Thorndike, E. (1904) *An Introduction to the Theory of Mental and Social Measurements*, Science Press.

おしお・あつし 早稲田大学文学学術院教授。著書に『性格とは何か——より良く生きるための心理学』（中央公論新社、2020年）など多数。パーソナリティ心理学、発達心理学専攻。