

# 職場での選抜研究

## ——面接、適性検査、エントリーシートに基づく選抜手法の理論と実証

鈴木 智之

(名古屋大学大学院准教授)

本稿は、企業が直接雇用する労働者を外部労働市場から募集・選考し、労働契約を結ぶという採用活動のうち、選抜研究の系譜に則って構成された。就職選抜手法によって付される選抜成績と入職後のジョブ・パフォーマンスの二地点間関係構造を仮定し、その上で新規学卒者選抜に焦点化して、面接、適性検査、エントリーシートの選抜妥当性研究の理論的背景と主な実証研究を論じることを目的にした。第一に、面接の理論的背景として選抜妥当性研究の礎となる信頼性、構造化、選抜基準となる構成概念を主に取りあげた。実証研究として1980年代以降になされた主要なメタ分析の世界的潮流を系統化して述べ、ジョブ・パフォーマンスの妥当性とその要因を論じた。第二に、適性検査の理論的背景として知的能力研究とパーソナリティ特性研究を取りあげた。実証研究として一般知的能力とビッグファイブ・パーソナリティに関する主要なメタ分析の潮流を述べ、ジョブ・パフォーマンスの妥当性とその要因を論じた。第三に、エントリーシートの理論的背景が薄弱であり、さらに実証研究が乏しく、僅かに存在する実証研究であっても理論的布置が今後の課題であることを論じた。

### 目次

- I 課題関心
- II 理論的布置と目的
- III 面接に関する理論研究の歴史・系譜と実証研究
- IV 適性検査に関する理論研究の歴史・系譜と実証研究
- V エントリーシートに関する理論研究の歴史・系譜と実証研究

### I 課題関心

採用とは、企業が直接雇用する労働者を外部労働市場から募集・選考し、労働契約を結ぶことである(今野・佐藤 2020)。そのうち、募集面に着目して、人材の誘引を論じる採用研究(recruitment research)と選考面に着目して、人材の合否決定法を論じる選抜研究(selection research)を下位

研究に設けることが一般的である(Wanous 1992など)。本稿は選抜研究の系譜に則る。

選抜は対象者の能力や特性などを評価する手法(以下、選抜手法)によってなされる。筆記試験、ワークサンプル・テスト、レファレンス・チェックなどの多様な選抜手法の中で、国内外の実践面での歴史を通して中心的位置を占めてきたのは面接である(Arvey and Campion 1982; Cortina et al. 2000; 荻谷・本田 2010)。ただし、その実務的普及が選抜手法としての適切性をそのまま意味するわけではないことにはかつてより注意が払われていた(Wiesner and Cronshaw 1988)。欧米の学術研究ではジョブ・パフォーマンスに最も妥当性を有する選抜手法は知的能力の筆記試験であり、それに比して面接の妥当性は大きく劣るという見方が少なくとも1980年代半ばまで支配的であり、そ

の頃まで実は面接は科学研究の主たる関心にすらなり得なかった (Hunter and Hunter 1984)。

面接を含む選抜手法の歴史を振り返ると、その適切性が複数の視角から論じられてきたことがわかる。例えば、公正研究による視角として Bies and Moag (1986) があり、人材選抜場面においてリクルータが求職者と意思疎通を図る際に敬意を示すこと、性別などに関連した質問を行わないこと (propriety of questions) が求職者の公正認知に影響を与えることが示された。これは公正を選抜手法の適切性の要件とした研究視角である。

選抜手法の適切性についての一般的な視角は就職試験と入職後の二地点を設けるものである。就職試験の選抜手法によって付された成績 (以下、選抜成績) と入職後のジョブ・パフォーマンスの二変数間関係を妥当性係数として報告するものだ。Schmidt and Hunter (1998) は、その二地点間関係構造に基づいて選抜成績とジョブ・パフォーマンスの関係性が全く認められない場合には、どの候補者を合格・入職させても差異はなく、選抜手法の存在意義が全くないと述べた。欧米では同構造を前提にして学術研究が豊富に蓄積されてきた。同構造に基づく選抜手法の妥当性研究を以下、選抜妥当性研究と呼ぶ。

広闊にこの構造を捉えるとき、School to Work Transition 研究 (Savickas 1999。以下、トランジション研究) にまで射程を伸長することもできる。入職前と入職後という二地点間関係構造に則り、その地点間の移行・変容のあり方を論じるのがトランジション研究の課題関心の特徴である。欧米だけでなく、わが国でも研究例が蓄積されてきた分野であるが、本来、移行・変容を論じるには追跡的アプローチが求められるのに対して、わが国の研究では縦断調査が少ないことが懸念される。入職後の職業人地点から入職前の大学生地点を振り返る方法 (回顧法) による横断的アプローチのトランジション研究には限界があることが指摘されている (中原・溝上 2014 など)。今後、トランジション研究を通して二地点間の移行・変容の力動が示されれば、選抜妥当性研究によって導出される妥当性係数の背景の一部が示唆されるため、わが国でも縦断的な研究が待ち望まれる。

日本企業を対象にした選抜妥当性研究例は非常に少ない。恣意的な二地点間比較にしか過ぎない分析例が散見されるなど、わが国では本格的議論に乏しい。今後の研究蓄積を見据えて、わが国の学術研究および企業実務が理解すべき欧米の選抜妥当性研究の特徴には大きく3点あるだろう。

第一に、選抜手法の信頼性 (reliability) である。一地点の変数が不安定な場合、二地点間の妥当性係数の解釈ができない。そのため、欧米では選抜妥当性研究の前提として信頼性についての実証分析が多数報告されるに至っている (例えば Latham et al. 1980 ; Conway, Jako and Goodman 1995)。

第二に、通状況の一貫性とその限界である。経団連による大規模調査 (日本経済団体連合会 2018) によると、日本企業の新規学卒者選抜基準の上位に「協調性」や「誠実性」などのパーソナリティ特性 (personality trait) があり、選抜基準にパーソナリティ特性が用いられる点において海外の動向 (Cortina et al. 2000 など) と共通している。パーソナリティ特性は適性検査や面接における選抜基準になる。

パーソナリティ特性には通状況の一貫性があるとされるが、それは万能なものではなく限界があることも広く指摘されてきた (例えば Mischel 1968 ; Mischel and Shoda 1995 ; Tett and Burnett 2003)。1960年代からの大論争 (いわゆる「人間-状況論争」) を経て、パーソナリティ特性はどのような状況でも一貫した行動発現に繋がるものではなく、状況による差異があることが指摘されるに至った。よって、誠実性や協調性などのパーソナリティ特性に関する選抜成績が高ければ、どのような状況 (ジョブや組織など) でもジョブ・パフォーマンスを創出する規定因になるとは言い切れず、一律的判断は危険性を伴う。状況要因を考慮した論証が必要になる。

第三に、継時的安定性と教育訓練可能性である。経団連による同調査では「ストレス耐性」が選抜基準の上位にあった。ストレス耐性と密接に関連する概念にレジリエンスがあり、精神的ストレスによる落ち込みから回復する力を指す<sup>1)</sup>。レジリエンスには先天的な因子と後天的に獲得可能な因子があり、職場場面でも社会人への教育訓練

によって高められることが実証的に示された (Robertson et al. 2015)。

二地点間関係構造には就職試験の選抜対象となる能力や特性が容易に変化しないという仮定がある。例えば、求職者のレジリエンスの高低が就職試験後すぐに容易に変わり得るとき、ジョブ・パフォーマンスとの関係を妥当性係数として算出する意義が低くなる。能力や特性が就職選抜時点から入職後にかけて容易に変化するような、教育訓練可能性が高い場合においては入職前の選抜基準というより、入職後の教育訓練対象として位置付けることができ、少なくとも静的な二地点間関係構造の論証には適さない。端的に言えば入職後に容易に高められる能力・特性は入職した後に教育訓練によって高めればよいのだから、入職可否の選抜基準として用いるべきではないということだ。

以上から一般的に、選抜手法の信頼性が高く、状況要因が考慮され、継時的安定性の高い能力・特性が選抜基準に用いられた上でジョブ・パフォーマンスとの高い妥当性係数が確認されたときに選抜手法の適切性が獲得されると解される。なお、公正研究の視角によれば適切性の定義が異なることになる。加えて、実務的な視角による適切性の定義もある。例えばアフーマティブ・アクションに基づいて特定の性別の者を多く新規採用するような実務場面においては採用目標人数の達成が第一の適切性として意識される。もちろんジョブ・パフォーマンスと無関係ではないが、企業の社会的責任などの二地点間関係構造とは別の視角から選抜手法としての適切性が複合的に検討されることになる。

## II 理論的布置と目的

以上の多様な視角を踏まえた上で本稿は、公正研究やアフーマティブ・アクションなどではなく、選抜研究の一般的な理論系譜に則って、就職選抜手法によって付される選抜成績と入職後のジョブ・パフォーマンスの二地点間関係構造を仮定し、その上で新規学卒者選抜に焦点化して、面接、適性検査、エントリーシートの選抜妥当性研

究の理論的背景と主な実証研究を論じることを目的にする。

同目的を達成する上では選抜妥当性研究の前提としての信頼性研究にまで射程を拡げる。さらに、単なる二地点間関係の論証ではなく状況および継時的安定性を考慮した検討を行う。これらに基づき、わが国では知見が十分蓄積されていない選抜妥当性研究についての理論研究の歴史・系譜と実証研究を体系的に論じることを本稿の新たな試みとする。

## III 面接に関する理論研究の歴史・系譜と実証研究

面接とは、人と人が一定の環境にあって、直接顔をあわせ、一定の目的を持って互いに話し合い、情報を交換したり、意志や感情を伝えたり、相談したり、問題を解決することを指し、人材雇用などの選抜を目的とするものが面接の最も古い利用形態とされる (保坂・中澤・大野木 2000)。

人間の長い歴史の中で面接が選抜手法として古来用いられてきた。紙資源が貴重だった時代にはイエズス会の学校などの一部の例外的組織を除いて筆記試験を行うことができなかつたためだ。その後、紙資源の普及とともに19世紀頃から面接が批判的検討の対象になり、面接における個々の面接者の偏向性や恣意性が懸念されて、米国マサチューセッツ州教育委員会における試験改革によってその改善が図られた (Hogan 2007; 繁榊 2014)。この系譜は複数面接者による評価の一致とその要因を探究する研究分野へと繋がり、選抜手法としての面接の評価者間信頼性 (inter-rater reliability) 研究として定式化されるに至った。

面接の評価者間信頼性の実証研究は欧米を中心に豊富に蓄積されており、Latham et al. (1980) の初期的研究で .76~.87、その後の Conway, Jako and Goodman (1995) によるメタ分析で .70 などの評価者間信頼性係数が報告された。面接の選抜妥当性研究の礎はこの信頼性の高さについての度重なる実証にあることを廃亡してはならない。

面接の選抜妥当性には先述の通り懐疑的見方がかつての常識だった。面接選抜の歴史を振り返

れば、科学的議論に耐えうる最初のメタ分析は Reilly and Chao (1982) によるものと思われ、面接選抜の入職後のジョブ・パフォーマンスへの妥当性係数は .19 を示した。その2年後に報告された Hunter and Hunter (1984) による分析でもほぼ同水準の係数が報告された。筆記試験と比べて選抜妥当性が低く、面接は選抜手法としての適切性を十分に獲得できるものではないと評価されていたことが当時の複数の文献で確認できる。

大きな転換点になったのは1980年代後半から実施された複数のメタ分析だった。Wiesner and Cronshaw (1988) によるメタ分析は面接の構造化と面接者の人数に着目した。その発見には大きく2つあり、非構造化から構造化することで面接成績のジョブ・パフォーマンスへの選抜妥当性が飛躍的に向上すること、非構造化面接では面接者を1名ではなく複数名配置することで選抜妥当性が向上することであった。面接の構造化によって高い選抜妥当性が得られることは、その後の Wright, Lichtenfels and Pursell (1989) のメタ分析でも確かめられた。そして1990年代になると構造化の有無の議論を脱してその程度に議論が進展し、Huffcutt and Arthur (1994) のメタ分析では半構造化面接よりも構造化面接のほうが高い選抜妥当性が得られることが実証的に報告された。

この頃には、面接の構造化についての議論にある程度の着地が見られ始めたことから構造化面接を前提にした上で状況要因の検討へと研究関心が移行した。1990年代後半になされた Schmidt and Rader (1999) によるメタ分析では、営業職と生産ライン技能職というジョブ別の面接の選抜妥当性が報告された。生産ライン技能職のほうが営業職よりも高い選抜妥当性を示し、その背景には業務の定型性と面接の構造化の関係がうかがえた。すなわち、生産ラインのように定型性の高い業務については質問内容や評定尺度をいわば定型化する構造化面接を用いることでジョブ・パ

フォーマンスを高い精度で予測できることが示唆された。ジョブに関連して、McDaniel et al. (1994) によって1990年代半ばになされたメタ分析では幅広く求職者の心理的特性を把握するような面接の質問よりもジョブに関連した質問を面接で問うことのほうが高い選抜妥当性を有することが報告された。

選抜手法としての面接の理論と実証に大きな影響を与えたのは選抜基準としての構成概念の理論化である。面接選抜成績として「合格・不合格」という総合結果だけではなく、面接で評価される求職者の能力・特性の細分化がなされたことで選抜妥当性を生む要因についての要素分解的議論が可能になった。構成概念と面接選抜の関係への関心の端緒は1910年代にまで遡ることができ、米国企業において一般知的能力 (general intelligence) とパーソナリティの面接への実装と検証がなされた (Lynch 1968)。

構成概念の理論化について本格的研究が進展したのは1990年代である。Schmidt and Hunter (1998) によるメタ分析では面接選抜基準としての一般知的能力に着眼された。表1の通り、一般知的能力を選抜基準に追加することで構造化面接では24%、非構造化面接では8%の選抜妥当性が増加した。すなわち、面接の選抜妥当性は構造化、面接者の配置人数、ジョブ状況、質問内容という先述した観点だけではなく、特定の構成概念に着眼することで向上することが実証的に示された。Cortina et al. (2000) によるメタ分析では知的能力だけでなくパーソナリティ特性 (誠実性) を面接選抜基準にしたときに高い選抜妥当性が得られることが実証的に報告された。

Schmidt and Hunter (1998) のもう1つの特徴は他の選抜手法の妥当性と比較した点だ。ジョブ・パフォーマンスへの妥当性係数としてワークサンプル・テストで .54、レファレンス・チェックで .26 を示し、構造化面接の選抜妥当性が相対

表1 面接の実証的な妥当性研究例

	妥当性 ( <i>r</i> )	妥当性の増分
構造化面接 (structured interview)	.51	24%
非構造化面接 (unstructured interview)	.38	8%

出所: Schmidt and Hunter (1998) をもとに筆者により抜粋して作表

的に高い水準にあることが報告された。ワークサンプル・テストとは入職後の実際の職務に類似した課題への取り組みを選抜に用いるものである。実際の職務課題に近い選抜手法と遜色ない選抜妥当性が面接の設計次第で得られるということだ。

これらの理論と実証の蓄積が面接へのかつての懐疑的見解を払拭し、面接の高い選抜妥当性とそれによる選抜手法としての適切性の獲得に繋がって、今日の選抜手法としての面接を支えている。

そのような潮流を経て、今日の海外の先進企業では一般知的能力やパーソナリティ特性を選抜基準にした構造化面接を導入している（Googleなど）。一方で、わが国の企業では構造化面接の導入こそ進んだが、信頼性や妥当性についてはあまり実証的に検証・報告されていないと推察され、選抜手法としての面接が企業内外でその適切性を十分に獲得・訴求し得る現状かどうかが懸念される。また、先述した経団連調査の選抜基準にはさまざまな概念が混在しており、構成概念の理論的・体系的知見が面接選抜基準に十分反映されていない可能性も危惧される。

昨今、オンライン面接が普及している。Blacksmith, Willford and Behrend (2016) によるメタ分析では、オンライン面接は対面面接に比して面接者の選抜成績が平均的に低くなり、求職者からの印象も平均的に悪くなることが示唆された。また、Gorman, Robinson and Gamble (2018) は動画面接を取りあげ、動画面接による選抜成績のジョブ・パフォーマンスへの選抜妥当性が正の方向にあることを報告したが、同研究では選抜以外の人物評価研究や多様な情報技術要素が混在し、サンプルサイズが小さいなどの研究法上の問題が見られ、面接のオンライン化に伴う影響についてはまだ不明な点が多い。

実務的普及が先行し、選抜手法としての適切性の検討がいわば後回しになっているという昨今のオンライン面接と動画面接の動向は1980年代半ばまでの対面面接の議論を想起させるものだ。かつての対面面接の論点が40年ほどしてオンライン面接と動画面接の論点に形を変えて眼前に浮上しているように思える。対面面接は長い間、選抜妥当性が低く、選抜手法としての適切性を獲得し

ないまま実務で普及し、組織と個人の不適合などを生じさせる一因になっていた。オンライン面接や動画面接もその轍を踏まぬよう、研究の歴史に倣うべきである。その参照点になる主な理論と実証は本稿で述べたものである。

## IV 適性検査に関する理論研究の歴史・系譜と実証研究

選抜手法としての適性検査は筆記試験によって求職者の能力や特性などを評価するもので、一般的に能力検査と性格検査によって構成される（高橋 2021）。

適性検査において能力と性格を分離して評価することは Ackerman and Heggestad (1997) によるメタ分析などに根拠を求められる。同研究では一般知的能力とビッグファイブ・パーソナリティの協調性および誠実性との相関係数が非有意であり、別個の独立した概念であることが示された。記憶や認知処理速度という他の知的能力の側面もそれらのパーソナリティ特性を含むビッグファイブとはほぼ無関連であることが示された。能力と性格それぞれの適性検査が選抜手法として必要であるということだ。

### 1 能力検査

能力検査は知的能力研究にその系譜を求めることができる。知的能力研究の歴史を振り返ると Spearman による g 因子の発見（Spearman 1904）、1905年の Binet と Simon による知能検査開発などにその研究端緒を見出せる。実践的見地に立つと、知的能力を用いた就職選抜の起源は中国で行われていた科挙にある。科挙は紀元前 200 年前後から開始されたとする文献もあり、西洋の筆記試験にも影響を与えた可能性がある（Hogan 2007）。

知的能力研究には大きく2つの立場がある。1つは一般知的能力の存在を前提にし、同知能の働きを強調する知能の2因子説（Carroll 1993 など）、もう1つは一般知的能力の存在を前提にしない知能の多因子説（Thurstone 1938 など）や Multiple Intelligence 理論（ガードナー 2001 など）である。例えば Multiple Intelligence 理論は一般知的能力

という共通因子は存在せず、言語、論理数学、対人など相互に独立した知的能力が存在すると規定する。

Schmidt and Hunter (1998) によるメタ分析では選抜基準として一般知的能力の筆記試験を用いたとき、入職後のジョブ・パフォーマンスへの妥当性係数は .51 であることが示された。選抜手法として最もジョブ・パフォーマンスへの妥当性を高く有するのは一般知的能力の筆記試験であると米国で伝統的に長い間認識されており、多数の研究によってその裏付けがなされてきたことで選抜手法としての適切性が獲得されてきた。なお、選抜手法としての面接にとって 1980 年代から 1990 年代にかけての最大の論点の 1 つは、この一般知的能力の筆記試験の選抜妥当性の水準に迫ることができるのかどうかだった。他の選抜手法の適切性獲得の基準値となるぐらい、一般知的能力の筆記試験の選抜妥当性が優れていたということだ。

一般知的能力の約 8 割は遺伝的要因によって規定されるが、言語に関する知的能力は約 1 割しか遺伝的要因によって規定されず、あとは環境的要因によること (太田 2019)、また、知的能力の中には流動性知能 (Gf) のように成人期になってからは発達しないものと結晶性知能 (Gc) のように成人期になってからも発達し続けるものがあることが知られている (Horn 1970)。つまり、知的能力の中には職場での教育訓練可能性が低く、継時的安定性が高いものと教育訓練可能性が高いものがあるということだ。少なくともわが国の面接や適性検査ではこの区分がなされているか不明であり、本来は入職後の教育訓練対象とすべき、いわゆる「頭の良さ」を入職前の就職選抜基準にしてしまっただけで能力検査として実装していないかが懸念される。

わが国の研究を対象にメタ分析を行った高橋・西田 (1994) は一般知的能力の職務業績評価への妥当性係数を .249 または .300 と報告したが、その後の国内研究ではそれよりかなり低い妥当性係数が報告された (飯塚ほか 2005) など、一貫した知見が導出されるには至っていない。その背景には知的能力を用いた能力検査について継時的安定性と教育訓練可能性が十分に検討されていないこ

と、知的能力の多様な各構成要素と個別のジョブの関係性が十分に実証されていないことがあるのではないかと考える。わが国における能力検査の適切性が実証されるにはそれらの論点を踏まえた研究展開が求められ、今後期待される。

実務実態として用いられる能力検査の多くは商用尺度であり、社外に知的能力の構成要素別の継時的安定性、教育訓練可能性、ジョブごとの選抜妥当性が公開されることはほぼなく、外部からは不明な点が多い。それが日本企業における理論と実証の両面での能力検査の適切性の乏しさや実務的な納得感の低さの要因になっている。

## 2 性格検査

性格検査はパーソナリティ研究にその系譜を求めることができ、歴史を遡れば 19 世紀後半になされた Galton (1884) などによる初期研究にその端緒を見出せる。また、20 世紀に開発された職場に関連する主要なパーソナリティ尺度の多くが他記式ではなく自記式であったため、選抜手法としての性格検査でもアセッサーなどの他者が評価するものではなく、求職者自身が認識する性格を回答するという今日の形態が用いられているものと推察される。性格検査の理論的背景には大きく 2 つの潮流がある。

第一に、継時的安定性である。パーソナリティ特性の継時的安定性を示す研究例は数多く、1932 年から 1994 年までになされた縦断調査に基づく 152 個の研究を用いたメタ分析ではパーソナリティ特性について成人後の安定性の高さが示された (Roberts and DelVecchio 2000)。さらに遺伝的要因 (McCrae et al. 2000) と後天的要因 (Roberts, Wood and Smith 2005) によって継時的安定性が認められることも報告された。これらによって選抜基準としての要件を満たすことが確認された。なお、レジリエンスの例のように一部個別のパーソナリティ特性については教育訓練による可変性が認められたものも存在するが、全般的にはパーソナリティ特性には高い継時的安定性が認められてきた。

第二に、状況要因を踏まえた選抜妥当性である。パーソナリティ特性理論を代表するものに

ビッグファイブ理論があり、1930年代の研究 (Allport and Odbert 1936) にまでその起源を遡ることができる。ビッグファイブとは人間のパーソナリティ特性を外向性、協調性、誠実性、情緒安定性、開放性という5個の因子で記述する理論である (例えば Goldberg 1990, 1992)。さまざまな国で信頼性と妥当性が確かめられ、極めて多数の研究によって確固たる知見を積み重ねている。海外だけではなく1990年代から現在に至るまで日本語版のビッグファイブ尺度も複数開発された。これらの理論と実証の蓄積が性格検査の選抜手法としての適切性を示す有力な根拠になっている。

選抜妥当性研究として世界的に参照される Barrick and Mount (1991) によるメタ分析では人材選抜場面におけるビッグファイブのジョブ・パフォーマンスへの妥当性として特に誠実性が高い選抜妥当性を有することが報告された。ビッグファイブとジョブ・パフォーマンスの妥当性は欧州を対象にした Salgado (1997) によるメタ分析などでも確認され、北米のみならず欧州における性格検査の適切性の根拠を提供した。

Barrick and Mount (1991) の特徴は単に5因子の選抜妥当性を報告するのみならず、ジョブの種類を分けて選抜妥当性の高低を実証的に示した点だ。主な分析結果を表2に示した。同じパーソナリティ特性であってもジョブによって妥当性係数が正と負を示すことがあり、状況要因の重要性が示された。

例えば、「協調性」は専門職や営業職のジョブではほぼ選抜妥当性を有しないことが同研究で示された。日本企業の新卒選抜基準としての適切性に疑問が持たれる結果だった。「情緒安定性」は日本企業の新卒選抜基準のうち「ストレス耐性」と関連する概念であるが、表2に示された通り、ジョブによって正負の妥当性係数が混在した。日本企業の実務実態としてストレス耐性が低ければジョブの詳細を深く考えずに一律に採用を慎重に考え、ときに採用を見送るような性格検査の運用がなされていることが少なくない。しかし、メタ分析によればジョブによって本来判断を変えるべきであるということだ。

2000年代になると、Barrick and Mount (1991)

表2 パーソナリティ特性の実証的な妥当性研究例

	分析対象者計	$\hat{\rho}$
外向性		
専門職	476	-.09
管理職	11,335	.18
営業職	2,316	.15
協調性		
専門職	557	.02
管理職	8,597	.10
営業職	2,344	.00
誠実性		
専門職	767	.20
管理職	10,058	.22
営業職	2,263	.23
情緒安定性		
専門職	518	-.13
管理職	10,324	.08
営業職	2,486	.07
開放性		
専門職	476	-.08
管理職	7,611	.08
営業職	1,566	-.02

出所：Barrick and Mount (1991) をもとに筆者作表

を含む、1990年代になされた15個のメタ分析についてのセカンドオーダー・メタ分析が Barrick, Mount and Judge (2001) によって行われ、状況要因の重要性がさらに強調された。主に、チームワークの形成という個別的状況に特化してジョブ・パフォーマンスを捉えると「情緒安定性」は  $\rho=.22$ 、上司による評価という個別的状況に特化してジョブ・パフォーマンスを捉えると「誠実性」は  $\rho=.31$  を示し、Barrick and Mount (1991) の妥当性係数よりも高い水準を示すことが報告された。その後になされた Hogan and Holland (2003) によるメタ分析では状況要因を考慮することでパーソナリティ特性のジョブ・パフォーマンスへの妥当性係数が最大で.34～.43まで高く得られることが実証的に報告された。

ジョブや求める成果の種類を適切に考慮すればパーソナリティ特性を測定する性格検査に高い選抜妥当性が期待できるということである。ストレス耐性に限らず、性格検査における状況的差異を無視した一律的な合否判断は世界的潮流と逆行する運用であり、性格検査の長年の研究で蓄積された妥当性の根拠を毀損することになりかねない。

## V エントリーシートに関する理論研究 の歴史・系譜と実証研究

エントリーシートとは、就職試験において企業から求職者に課される質問形式の応募書類を指す。自己PRや志望動機についての質問項目が主で、課題論述が含まれることもある（上林・厨子・森田 2010 など）。また、「学生時代に力を入れたこと（実務的にはガクチカと略称される質問項目）」が含まれることも少なくない。

この選手法が面接や適性検査と大きく異なるのは歴史・系譜を伴う理論的背景ではなく実務的要請に導かれて普及したという点だ。面接は信頼性と選抜妥当性の研究蓄積を根拠に、また適性検査は知的能力とパーソナリティの各領域の研究蓄積を根拠にして選手法としての適切性を獲得してきた。しかし、エントリーシートは一時期一斉採用によって求職者全員を面接することができないことへの実務的対応から導出されたものだ。

広く文章評価や大学入試の小論文評価にまで射程を拡張すれば理論的背景の一部を探索することはできるが、実務実態としてそれらに根付いた選抜がなされているわけではなく、研究知見も現状分散的で収束していない。例えば、高校生が特定の論述課題について自身の認知的側面を働かせて記述する大学入試の小論文と大学生が自己PRや志望動機などのいわば情意面を多く含む非認知的側面を働かせて記述するエントリーシートを同等に扱ってよいと考えるのは仮定が強く外的妥当性上の懸念が残るため、両分野の越境的知見が今後提示されるべきだろう。エントリーシートに関する科学的な研究知見はほぼ見られず、選抜場面の学術研究は皆無であるという報告もある（鷲坂・二村・山岸 2001）。

この理論的背景の薄弱さは昨今の人口知能の活用動向にも見てとれる。大量のエントリーシートの自動評価を目指して人口知能を導入する会社がわが国でも見られ始めているが、人工知能がエントリーシートの「何をどう判断しているのかについては完全にブラックボックスであり、私たちも分からない」という日本企業が存在する（ITメ

ディア 2017）。選抜基準が不明なままでは妥当性係数が仮に算出されてもその意味内容への理解が及ばない。人材採用投資への説明責任を果たすことができず、求める人物像の定義と発信もできないため、選抜面だけでなく募集・誘引面にも悪影響を及ぼしかねず、危惧される状況である。

人工知能技術を用いる前に、本来はエントリーシートの文章から求職者のどのような能力・特性が捕捉可能か、その限界は何か明らかにされるべきである。外的妥当性に留意しながらも、文章の書き手の非認知的側面を把握した研究に立ち戻る必要がある。文章と非認知的側面の関係について取り組んだ、世界で最初の科学研究はオールポート（1982）によるものだろう。1926年から1937年までの11年半にやり取りされた手紙から書き手の特性を探索するという意欲的な研究であり、文章だけから書き手の攻撃性、自己中心性、情緒不安定性、根気強さのようなパーソナリティ特性が抽出された。原典は1946年にG. W. Allportが発表した研究成果であり、欧米の心理学分野では非常に有名な基礎研究として古典的テキストになっている。わが国のエントリーシート選抜に関わる研究・実務ではこのようないわば教科書的な研究すらほぼ知られておらず、今後諸研究分野の知見を総合してエントリーシートの理論を定めなければならない。

エントリーシートに関する実証的な選抜妥当性研究は極めて希少だが、鈴木（2016）によるものがあり、面接対象者を絞り込むための予備選抜としてのエントリーシートの特徴に着眼して、基準変数にエントリーシート選抜のすぐ後に実施される一次面接の合否を用いた妥当性研究である。同研究は日本企業の新規学卒者就職試験で実際に学生から提出されたエントリーシートを取得し、自己PRや志望動機などの自由記述文について形態素解析と頻度分析を行って予測変数とし、基準変数の予測判別を試みた。その結果、エントリーシートに記述された3個の語（「達成」「リーダー」「イベント」）の出現頻度から一次面接成績を約72%の正判別率で予測できることが示された（表3）。なお、表3の予測値はエントリーシートの3語による一次面接成績の予測結果、実績値は実際

表3 エントリーシートの実証的な妥当性研究例

	実績値		計
	一次面接合格群	一次面接不合格群	
予測値	一次面接合格群	7	25
	一次面接不合格群	15	21
	計	22	46

出所：鈴木（2016）

の各求職者の一次面接での成績を指した。

達成やリーダーの語については達成動機理論やリーダーシップ理論から解釈することも可能だが、恣意的な理論適用にならないように留意が必要だ。探索的な研究知見の蓄積は重要だが、理論的枠組みに布置された研究知見が蓄積されることによってエントリーシートの妥当性が解釈され、選抜手法としての適切性が獲得される。また、二地点間関係構造に基づいて面接や適性検査と同様にジョブ・パフォーマンスが基準変数に設けられなければならないが、面接成績までの射程にとどまっている点も今後の研究課題である。

以上の通り、現状、エントリーシートは理論・実証共に不足したまま実務実態のみが先走っており、選抜手法としての適切性が十分担保されているとは言えない。今後の理論検討と実証分析の蓄積が強く求められる。

謝辞 本論文執筆の文献収集にあたってJSPS 科研費JP22K01697の助成を受けた。

1) レジリエンスは能力か特性かという議論があるが、本稿では簡便な記載にとどめた。

#### 参考文献

ITメディア（2017）「HR Tech 最前線 ソフトバンクが新卒の「ES 選考」を AI に任せた理由」。<https://www.itmedia.co.jp/business/articles/1708/29/news011.html>（閲覧日：2022 年 7 月 17 日）

飯塚彩・持主弓子・内藤淳・二村英幸（2005）「適性検査の予測的妥当性——職種別および製造・非製造業別の分析」『産業・組織心理学会第 21 回大会発表論文集』pp. 159-162.

今野浩一郎・佐藤博樹（2020）『マネジメント・テキスト 人事管理入門（第 3 版）』日本経済新聞出版。

オールポート W. G. / 青木孝悦・萩原滋訳（1982）『ジェニーからの手紙——心理学は彼女をどう解釈するか』新曜社。

太田信夫監修（2019）『シリーズ心理学と仕事 9 知能・性格心理学』北大路書房。

ガードナー、ハワード / 松村暢隆訳（2001）『MI——個性を生かす多重知能の理論』新曜社。

菊岡剛彦・本田由紀（2010）『大卒就職の社会学——データからみる変化』東京大学出版会。

上林憲雄・厨子直之・森田雅也（2010）『経験から学ぶ人的資源管理』有斐閣。

Google「構造化面接を実施する」。<https://rework.withgoogle.com/jp/guides/hiring-use-structured-interviewing#introduction>（閲覧日：2022 年 12 月 26 日）

鷺坂由紀子・二村英幸・山岸建太郎（2001）「採用選考における作文評価——達成動機測定を試み」『経営行動科学』第 14 巻第 3 号, pp. 153-159.

繁榊算男（2014）『新しい時代の大学入試』金子書房。

鈴木智之（2016）「新規学卒者採用試験における選考書類内の形態素と採用面接成績との関連についての実証分析」『日本労務学会誌』17 巻 1 号, pp. 19-35.

高橋潔（2021）「形のない経済（インタンジブル・エコノミー）における採用・評価・育成——産業・組織心理学の視点」『産業・組織心理学研究』35 巻 1 号, pp. 3-23.

高橋潔・西田直史（1994）「知的能力検査に関する妥当性一般化——メタ分析による結果」『産業・組織心理学研究』8 巻 1 号, pp. 3-12.

中原淳・溝上慎一編（2014）『活躍する組織人の探究——大学から企業へのトランジション』東京大学出版会。

日本経済団体連合会（2018）「2018 年度新卒採用に関するアンケート調査結果」。

保坂亨・中澤潤・大野木裕明（2000）『心理学マニュアル——面接法』北大路書房。

Ackerman, P. L. and Heggestad, E. D. (1997) "Intelligence, Personality, and Interests: Evidence for Overlapping Traits," *Psychological Bulletin*, Vol. 121, No. 2, pp. 219-245.

Allport, G. W. and Odbert, H. S. (1936) "Trait-Names: A Psycho-Lexical Study," *Psychological Monographs*, Vol. 47, No. 1, pp. i-171.

Arvey, R. D. and Campion, J. E. (1982) "The Employment Interview: A Summary and Review of Recent Research," *Personnel Psychology*, Vol. 35, No. 2, pp. 281-322.

Barrick, M. R. and Mount, M. K. (1991) "The Big Five Personality Dimensions and Job Performance: A Meta-Analysis," *Personnel Psychology*, Vol. 44, No. 1, pp. 1-26.

Barrick, M. R., Mount, M. K. and Judge, T. A. (2001) "Personality and Performance at the Beginning of the New Millennium: What Do We Know and Where Do We Go Next?" *International Journal of Selection and Assessment*, Vol. 9, No. 1-2, pp. 9-30.

Bies, R. J. and Moag, J. S. (1986) "Interactional Justice: Communication Criteria of Fairness," in R. J. Lewicki, B. H. Sheppard and M. H. Bazerman (eds.) *Research on Negotiation in Organizations Vol. 1*. pp. 43-55. Greenwich, CT: JAI Press.

Blacksmith, N., Willford, J. C. and Behrend, T. S. (2016) "Technology in the Employment Interview: A Meta-Analysis and Future Research Agenda," *Personnel Assessment and Decisions*, Vol. 2, No. 1, pp. 12-20.

Carroll, J. B. (1993) *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*, New York, NY: Cambridge University Press.

Conway, J. M., Jako, R. A. and Goodman, D. F. (1995) "A Meta-Analysis of Interrater and Internal Consistency Reliability of Selection Interviews," *Journal of Applied Psychology*, Vol. 80, No. 5, pp. 565-579.

Cortina, J. M., Goldstein, N. B., Payne, S. C., Davison, H. K. and Gilliland, S. W. (2000) "The Incremental Validity of

- Interview Scores over and above Cognitive Ability and Conscientiousness Scores," *Personnel Psychology*, Vol. 53, No. 2, pp. 325-351.
- Galton, F. (1884) "Measurement of Character," *Fortnightly Review*, Vol. 36, pp. 179-185.
- Goldberg, L. R. (1990) "An Alternative "Description of Personality": The Big-Five Factor Structure," *Journal of Personality and Social Psychology*, Vol. 59, No. 6, pp. 1216-1229.
- (1992) "The Development of Markers for the Big-Five Factor Structure," *Psychological Assessment*, Vol. 4, No. 1, pp. 26-42.
- Gorman, C. A., Robinson, J. and Gamble, J. S. (2018) "An Investigation into the Validity of Asynchronous Web-Based Video Employment-Interview Ratings," *Consulting Psychology Journal: Practice and Research*, Vol. 70, No. 2, pp. 129-146.
- Hogan, J. and Holland, B. (2003) "Using Theory to Evaluate Personality and Job-Performance Relations: A Socioanalytic Perspective," *Journal of Applied Psychology*, Vol. 88, No. 1, pp. 100-112.
- Hogan, T. P. (2007) *Psychological Testing: A Practical Introduction. Second Edition*. New York, NY: John Wiley and Sons.
- Horn, J. L. (1970) "Organization of Data on Life-Span Development of Human Abilities," in L. R. Goulet and P. B. Baltes (eds.) *Life-Span Developmental Psychology: Research and Theory*. pp. 423-466. New York, NY: Academic Press.
- Huffcutt, A. I. and Arthur, W. (1994) "Hunter and Hunter (1984) Revisited: Interview Validity for Entry-Level Jobs," *Journal of Applied Psychology*, Vol. 79, No. 2, pp. 184-190.
- Hunter, J. E. and Hunter, R. F. (1984) "Validity and Utility of Alternative Predictors of Job Performance," *Psychological Bulletin*, Vol. 96, No. 1, pp. 72-98.
- Latham, G. P., Saari, L. M., Pursell, E. D., and Campion, M. A. (1980) "The Situational Interview," *Journal of Applied Psychology*, Vol. 65, No. 4, pp. 422-427.
- Lynch, E. C. (1968) "Walter Dill Scott: Pioneer Industrial Psychologist," *Business History Review*, Vol. 42, No. 2, pp. 149-170.
- McCrae, R. R., Costa, P. T., Jr., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M. D., Sanz, J., Sánchez-Bernardos, M. L., Kusdil, M. E., Woodfield, R., Saunders, P. R. and Smith, P. B. (2000) "Nature Over Nurture: Temperament, Personality, and Life Span Development," *Journal of Personality and Social Psychology*, Vol. 78, No. 1, pp. 173-186.
- McDaniel, M. A., Whetzel, D. L., Schmidt, F. L. and Maurer, S. D. (1994) "The Validity of Employment Interviews: A Comprehensive Review and Meta-Analysis," *Journal of Applied Psychology*, Vol. 79, No. 4, pp. 599-616.
- Mischel, W. (1968) *Personality and Assessment*. New York, NY: John Wiley and Sons Inc.
- Mischel, W. and Shoda, Y. (1995) "A Cognitive-Affective System Theory of Personality: Reconceptualizing Situations, Dispositions, Dynamics, and Invariance in Personality Structure," *Psychological Review*, Vol. 102, No. 2, pp. 246-268.
- Reilly, R. R. and Chao, G. R. (1982) "Validity and Fairness of Some Alternative Employee Selection Procedures," *Personnel Psychology*, Vol. 35, No. 1, pp. 1-62.
- Roberts, B. W. and DelVecchio, W. F. (2000) "The Rank-Order Consistency of Personality Traits from Childhood to Old Age: A Quantitative Review of Longitudinal Studies," *Psychological Bulletin*, Vol. 126, No. 1, pp. 3-25.
- Roberts, B. W., Wood, D. and Smith, J. L. (2005) "Evaluating Five Factor Theory and Social Investment Perspectives on Personality Trait Development," *Journal of Research in Personality*, Vol. 39, No. 1, pp. 166-184.
- Robertson, I. T., Cooper, C. L., Sarkar, M. and Curran, T. (2015) "Resilience Training in the Workplace from 2003 to 2014: A Systematic Review," *Journal of Occupational and Organizational Psychology*, Vol. 88, No. 3, pp. 533-562.
- Salgado, J. F. (1997) "The Five Factor Model of Personality and Job Performance in the European Community," *Journal of Applied Psychology*, Vol. 82, No. 1, pp. 30-43.
- Savickas, M. L. (1999) "The Transition from School to Work: A Developmental Perspective," *Career Development Quarterly*, Vol. 47, No. 4, pp. 326-336.
- 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.
- Schmidt, F. L. and Rader, M. (1999) "Exploring the Boundary Conditions for Interview Validity: Meta-Analytic Validity Findings for a New Interview Type," *Personnel Psychology*, Vol. 52, No. 2, pp. 445-464.
- Spearman, C. (1904) "'General Intelligence,' Objectively Determined and Measured," *American Journal of Psychology*, Vol. 15, No. 2, pp. 201-292.
- 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.
- Thurstone, L. L. (1938) *Primary Mental Abilities*, Chicago, IL: The University of Chicago Press.
- Wanous, J. P. (1992) *Organizational Entry: Recruitment, Selection, Orientation, and Socialization of Newcomers second edition*, Reading, Mass: Addison-Wesley Pub.
- Wiesner, W. H. and Cronshaw, S. F. (1988) "A Meta-Analytic Investigation of the Impact of Interview Format and Degree of Structure on the Validity of the Employment Interview," *Journal of Occupational Psychology*, Vol. 61, No. 4, pp. 275-290.
- Wright, P. M., Lichtenfels, P. A. and Pursell, E. D. (1989) "The Structured Interview: Additional Studies and a Meta-Analysis," *Journal of Occupational Psychology*, Vol. 62, No. 3, pp. 191-199.

すずき・ともゆき 名古屋大学大学院経済学研究科准教授。主著に『ワークプレイス・パーソナリティ論——人的資源管理の新視点と実証』（東京大学出版会、2023年）。人的資源管理論専攻。