

適性検査の予測的妥当性

- 適性検査と職務遂行能力評価の収集時期のひらきからの分析 -

持主 弓子 今城 志保 内藤 淳 二村 英幸
(HRR 株式会社)

Predictive validity of personality and general mental ability tests:
Analysis by the timing of performance evaluation

Yumiko Mochinushi Shiho Imashiro Jun Naito Hideyuki Nimura
(HRR Inc.)

1. 問題の背景

適性検査は採用選考など様々な場面で利用されている。現在の経営人事においては意思決定のための有力な材料の1つであり、その妥当性を確認しておくことは重要なことと思われる。

米国においてはこれまで数多くの適性検査に関する妥当性研究が行われてきた。性格検査の妥当性の水準については、Barrick & Mount (1991) が主要性格5特性 (Big5) の枠組みを用いて性格検査に関する妥当性研究のメタ分析を行い、Conscientiousness (誠実性) がさまざまな職種平均で、.22 (補正值) の妥当性係数を示すことを報告している。一般知的能力検査については Hunter & Hunter (1984) が.53の妥当性係数を報告している。日本の妥当性研究においては職務遂行能力と性格検査、または一般知的能力検査との関連をみた妥当性研究は少ないのが現状であるものの、いくつかの尺度において注目すべき妥当性係数が報告されている (高橋・西田, 1994; 二村・今城・内藤, 2000; 大沢・芝・二村, 2000)。

一方で、これまでの妥当性研究と適性検査の実際の使用場面を照らし合わせてみたとき、考慮してみるべき点が2つ存在すると思われる。1つは妥当性研究の多くが予測変数と基準変数を同時期に収集する並存的妥当性を検討していることである。ところが実際の適性検査の利用状況を考えると、採用選考時に入社後の職務遂行能力を予測する参考資料として使用されることが多い。つまり、予測変数である適性検査と基準変数である職務遂行能力の評価には時間の隔たりがあることを前提にして使用しているといえる。故に採用時に適性検査を実施し、入社後に職務遂行能力を評価する予測的研究デザインのデータのみで妥当性係数を確認してみることが必要と思われる。

もう1つは、適性検査が実施された時期と基準

変数である職務遂行能力を評価した時期のひらきが妥当性係数に対し、どのような影響を与えるのかという点である。適性検査の使用者にとっては、検査が入社後のどの時点の職務遂行能力を予測するものかはひとつの関心事である。妥当性係数に影響をあたえるものとしてこの他に、職種や基準変数の性質等が挙げられるが、これらについては既に Barrick & Mount や Hunter & Hunter などの先行研究において考察が加えられている。しかし、基準変数である職務遂行能力の評価のタイミングをかえて予測変数との関係を確認した研究は Dreher & Bretz (1991) 以外は、日本での研究も含めてみあたらなかった。

本研究では性格検査と一般知的能力検査の2つにおいて、予測的研究デザインで行われた研究のみを使用してメタ分析の手法を用いて妥当性係数の真値の推定を行うとともに、検査実施と基準変数である職務遂行能力評価が行われた時点のひらきによって、妥当性係数の水準がどのように異なるのかを確認することを目的とする。

2. 分析方法

<対象企業>

総合検査 SPI (Synthetic Personality Inventory・HRR 株式会社) を採用選考時または内定者や新入社員の時期に実施した42社のデータを使用している。この42社は製造11社、非製造31社であり、非製造の中には金融、サービス、商社など様々な業種が含まれている。

<予測変数と基準変数>

予測変数に総合検査 SPI を用いた。実施は全て採用選考時または内定前後で行われている。この検査は性格検査17尺度および一般知的能力検査2尺度とその総合の計20尺度から構成されている。信頼性係数 (Cronbach の係数) は性格検査尺度で.80~.91、一般知的能力検査尺度で.83

～.89である。

基準変数は妥当性研究の目的で、人事考課や昇進スピードなどをもとに人事部門等で対象者の職務遂行能力を評価したものが主となっている。予測変数は1982年7月～2002年11月に、基準変数は1998年4月～2003年2月に収集された。

<対象研究>

始めに対象企業ごとに妥当性係数を求めた42個の妥当性研究を、メタ分析的手法を用いて再計算し妥当性係数の真値の推定を行った。各個別研究の平均サンプル数は274であった。

次に対象企業ごとに適性検査が実施された時期と職務遂行能力を評価した時期のひらきによって、5年まで、6～10年、11～15年の3グループに分け、それぞれ26、26、12、計64の妥当性係数を得た。各個別研究の平均サンプル数は176(32～833)であった。

<計算の手続き>

計算の手続きは、Hunter & Schmidt (1990)の妥当性一般化研究に従った。この手続きは、先行研究であるBarrick & Mount (1991)や高橋・西田(1994)と同じで、個別企業の妥当性係数を総合し、一般的水準を推定する。予測変数は標準化された検査であるため、信頼性係数及び範囲制限について把握することができ、その値をもって補正を行っている。一方、基準変数は人事考課や昇進スピードなどをもとに人事部門等で対象者の職務遂行能力を評価したものが主となっているため信頼性係数は算出できない。そこで先行研究(丸山・二村・内藤・今城, 2002)を参考にし、.60を用いて補正を行った。

3. 結果

表1は企業別に算出した42個の妥当性係数をメタ分析したものである。推定された妥当性係数の真値()は「社会的内向性」(.15)、「身体活動性」(.16)、「活動意欲」(.16)、「高揚性」(.11)、「内向-外向」(.14)、「非言語」(.12)で.10を超える値となった。

次に80%確信区間を確認する。80%確信区間とは、推定された妥当性係数の真値が収まる区間を検証するものであり、この区間に0を含まず、正もしくは負で一定水準の値となっていれば、意味のある値をとるものとする。上記の6尺度中「非言語」を除く5尺度が0を含まない値となっており、これらの性格尺度は異なる状況においても、

【表1】 42社でのメタ分析結果

	42社でのメタ分析結果							
	補正前平均妥当性係数	補正後の妥当性係数 ρ	妥当性係数の真値の分散	誤差による分散の明率	2	r ²	e ²	80%確信区間
社会的内向性	-0.10	-0.15	0.007	44.7%	0.007	0.013	0.006	-0.259 - -0.041
内省性	-0.04	-0.05	0.004	54.7%	0.004	0.010	0.005	-0.137 - 0.032
身体活動性	0.11	0.16	0.008	40.2%	0.008	0.013	0.005	0.044 - 0.268
持続性	-0.03	-0.03	0.003	64.2%	0.003	0.008	0.005	-0.100 - 0.033
慎重性	-0.04	-0.05	0.001	89.1%	0.001	0.005	0.005	-0.077 - -0.014
達成意欲	0.05	0.09	0.006	48.8%	0.006	0.012	0.006	-0.015 - 0.189
活動意欲	0.12	0.16	0.005	49.1%	0.005	0.010	0.005	0.068 - 0.255
敏感性	-0.03	-0.05	0.010	36.6%	0.010	0.016	0.006	-0.181 - 0.077
自責性	-0.04	-0.06	0.005	54.3%	0.005	0.011	0.006	-0.153 - 0.028
気分性	-0.02	-0.03	0.002	71.6%	0.002	0.009	0.006	-0.090 - 0.035
独自性	-0.05	-0.08	0.011	34.0%	0.011	0.017	0.006	-0.211 - 0.061
自信性	0.03	0.04	0.003	70.2%	0.003	0.008	0.006	-0.020 - 0.108
高揚性	0.07	0.11	0.006	48.9%	0.006	0.012	0.006	0.010 - 0.210
内向-外向	0.10	0.14	0.005	51.4%	0.005	0.010	0.005	0.050 - 0.225
直観-感覚	-0.01	-0.01	0.003	65.6%	0.003	0.008	0.005	-0.081 - 0.055
感情-思考	0.00	-0.01	0.003	60.6%	0.003	0.009	0.005	-0.081 - 0.067
知覚-判断	-0.02	-0.02	0.001	89.0%	0.001	0.006	0.005	-0.056 - 0.008
言語能力	0.02	0.03	0.004	61.2%	0.004	0.011	0.007	-0.054 - 0.115
非言語能力	0.08	0.12	0.017	27.0%	0.017	0.023	0.006	-0.049 - 0.280
一般的的能力総合	0.06	0.09	0.012	34.1%	0.012	0.018	0.006	-0.051 - 0.231

一定の方向性を示す妥当性係数であることが認められた。

次に42社のデータを予測変数と基準変数が収集された時期のひらきによって3グループに分け、妥当性係数を再計算しメタ分析を行った(表2)。5年までのグループでは推定された妥当性係数の真値()は「社会的内向性」(.12)、「身体活動性」(.11)、「達成意欲」(.14)、「活動意欲」(.14)、「内向-外向」(.12)の5つの尺度で.10を超える(または-.10より小さい)結果となった。またこれらの尺度の80%確信区間を確認すると、全て0を含まない値となっている。6～10年グループでは推定された妥当性係数の真値()では「社会的内向性」(.16)、「身体活動性」(.19)、「活動意欲」(.18)、「高揚性」(.12)、「内向-外向」(.17)、「非言語」(.20)の6つの尺度で.10(または-.10)を超え、80%確信区間は全て0を含まない値となっている。11～15年グループでは推定された妥当性係数の真値()では「社会的内向性」(.23)、「内省性」(.13)、「身体活動性」(.25)、「達成意欲」(.12)、「活動意欲」(.24)、「独自性」(.18)、「高揚性」(.11)、「内向-外向」(.19)、「非言語」(.13)の9つの尺度で.10(または-.10)を超え、「内省性」、「身体活動性」、「活動意欲」、「独自性」、「内向-外向」、「非言語」の6尺度で80%確信区間に0を含まない値となった。この3グループを比較すると予測変数と基準変数がとられた時期のひらきがあるグループほど妥当性係数が.10(または-.10)を超える尺度が増え、妥当性係数の値自体も大きくなっている。

4. 考察

本研究では、いくつかの性格検査尺度に関し

て.10 (または, -.10) を超える妥当性係数が確認されるとともに、予測変数と基準変数が収集された時期のひらきに応じて妥当性係数の値が変化することが確認された。一方、一般知的能力検査では性格検査と異なり、中間の6~10年のグループのみ妥当性係数が.15を超える結果となった。

表1の42個の妥当性係数をメタ分析した結果では、性格検査は先行研究(丸山ら, 2002)と似た結果を示し、予測的研究デザインのみメタ分析においても同水準の値が得られた。一般知的能力検査については先行研究である高橋・西田(1994)で示された.300、二村ら(2000)で示された.257と比較して低い値となっている。Hunter & Hunter (1984)の職務群別の研究では一般知的能力検査の妥当性係数は職務群により.23~.56(補正值、職務遂行能力評価が基準)のパラツキがあることを報告している。また高橋・西田(1994)でも職種によって統制を行った結果では.216~.352の値のパラツキがみられた。本研究では職種の統制は行っておらず、その構成は営業系(22.1%)、事務系(10.3%)、技術・開発系(22.4%)、その他(1.5%)、不明(43.6%)となっている。このような様々な職種で構成された標本であるため、職種特性の影響により妥当性係数が低い値にとどまった可能性がある。さらに二村ら(2000)は一般に、一般知的能力が求められる管理者を対象とした研究であり、高橋・西田(1994)についても職種別の結果は管理者のデータのみ用いられている。本研究では管理者以外のデータも多く含まれており、こうしたサンプルの違いが妥当性係数を低い値にとどめた可能性も考えられる。

また、二村ら(2000)は一般知的能力検査の妥

当性係数が.40~.50と報告されている米国の研究と比較して、日本においては本人の態度ややる気等を含めて職務遂行能力として評価する傾向が米国に比べて強く、その結果として一般知的能力検査の妥当性係数の水準が低めにとどまることを考察している。今回の一般知的能力検査に関する結果も同様の理由による可能性がある。

性格検査の妥当性係数の値が予測変数と基準変数のとられた時期のひらきに応じて異なった点については3つのポイントから考察を加えることができる。まず1つは、勤続年数による評価観点の違いである。「社会的内向性」「身体活動性」「活動意欲」「内向・外向」は予測変数と基準変数のとられた時期のひらきが大きくなると、妥当性係数の値が緩やかに上昇する傾向がみられた。これは新入社員のうちは協調的でフットワークのよい人物が評価されるが、後輩や部下ができるにつれて、これらに加えて社交性や実行力、周囲に配慮する姿勢、人脈を構築する力等も求められる結果、妥当性係数の水準が高まったと考えられる。また上記4尺度と異なり、「達成意欲」が上昇していないのは、階層が高くなるに従い意欲的で前向きな態度だけではなく、確実に実行できるか否かが問われるようになることから単純に妥当性係数の水準が高まるものではないと思われる。「高揚性」についても明るく楽天的なだけでなく、社交性も求められるようになることから同様に妥当性水準が大幅には高まらないと考えられる。さらに11~15年グループの「内省性」「独自性」で.10を超える妥当性係数が見られたのは、後輩・部下など周囲の意見にも耳を傾け吸収していく柔軟性が求められることを示唆しているのかもしれない。

【表2】適性検査を実施した時期と職務遂行能力を評価した時期のひらきごとのメタ分析結果

	5年まで					6~10年					11~15年				
	補正前平均妥当性係数	補正後の妥当性係数 ρ	妥当性係数の真値の分散	誤差による分散説明率	80%確信区間	補正前平均妥当性係数	補正後の妥当性係数 ρ	妥当性係数の真値の分散	誤差による分散説明率	80%確信区間	補正前平均妥当性係数	補正後の妥当性係数 ρ	妥当性係数の真値の分散	誤差による分散説明率	80%確信区間
社会的内向性	-0.08	-0.12	0.004	69.1%	-0.200 ~ -0.043	-0.11	-0.16	0.013	37.3%	-0.307 ~ -0.018	-0.14	-0.23	0.040	25.8%	-0.486 ~ -0.023
内省性	0.00	0.00	0.004	63.5%	-0.079 ~ -0.089	-0.05	-0.07	0.000	93.9%	-0.096 ~ -0.042	-0.09	-0.13	0.002	89.3%	-0.180 ~ -0.079
身体活動性	0.08	0.11	0.002	80.6%	0.055 ~ 0.163	0.13	0.19	0.009	44.1%	0.069 ~ 0.305	0.17	0.25	0.011	53.9%	0.114 ~ 0.386
持続性	-0.01	-0.01	0.003	67.4%	-0.081 ~ 0.066	-0.01	-0.02	0.001	91.1%	-0.050 ~ 0.014	-0.03	-0.04	0.014	47.0%	-0.190 ~ 0.118
慎重性	-0.02	-0.02	0.000	100.2%	-0.023 ~ -0.023	-0.04	-0.05	0.002	72.6%	-0.117 ~ 0.010	-0.03	-0.03	0.018	42.2%	-0.203 ~ 0.139
達成意欲	0.09	0.14	0.008	49.0%	0.026 ~ 0.261	0.05	0.09	0.007	53.1%	-0.020 ~ 0.198	0.07	0.12	0.012	57.8%	-0.017 ~ 0.258
活動意欲	0.10	0.14	0.000	104.3%	0.136 ~ 0.136	0.13	0.18	0.010	40.5%	0.049 ~ 0.302	0.16	0.24	0.011	55.1%	0.102 ~ 0.372
感受性	-0.03	-0.05	0.014	36.3%	-0.200 ~ 0.102	-0.03	-0.04	0.003	73.6%	-0.112 ~ 0.025	-0.06	-0.09	0.009	63.4%	-0.209 ~ 0.034
自費性	-0.05	-0.08	0.015	34.6%	-0.233 ~ 0.080	-0.04	-0.06	0.006	58.2%	-0.162 ~ 0.032	-0.03	-0.04	0.009	64.5%	-0.162 ~ 0.076
気分性	-0.01	-0.01	0.006	55.7%	-0.113 ~ 0.088	-0.01	-0.02	-0.001	107.6%	-0.023 ~ -0.023	-0.04	-0.08	0.012	58.5%	-0.220 ~ 0.055
独自性	0.00	0.00	0.010	43.4%	-0.130 ~ 0.130	-0.06	-0.09	0.006	58.1%	-0.185 ~ 0.007	-0.11	-0.18	0.000	100.9%	-0.177 ~ -0.177
自信性	0.05	0.07	0.006	58.5%	-0.027 ~ 0.170	0.03	0.05	0.004	66.1%	-0.033 ~ 0.129	0.02	0.04	-0.006	162.4%	0.040 ~ 0.040
高揚性	0.06	0.09	0.004	67.7%	0.008 ~ 0.167	0.08	0.12	0.009	45.7%	0.002 ~ 0.244	0.08	0.11	0.023	39.6%	-0.081 ~ 0.307
内向・外向	0.08	0.12	0.007	52.7%	0.009 ~ 0.224	0.13	0.17	0.010	40.9%	0.042 ~ 0.294	0.13	0.19	0.008	62.1%	0.072 ~ 0.302
直観・感覚	-0.01	-0.01	0.000	96.9%	-0.034 ~ 0.006	-0.01	-0.01	0.007	53.2%	-0.113 ~ 0.097	0.00	0.00	0.008	63.7%	-0.120 ~ 0.113
感情・思考	0.02	0.03	-0.001	109.9%	0.031 ~ 0.031	-0.01	-0.01	0.005	61.8%	-0.103 ~ 0.075	0.00	0.00	0.004	80.2%	-0.078 ~ -0.078
知覚・判断	0.02	0.02	0.000	94.3%	-0.007 ~ 0.048	-0.02	-0.02	-0.001	115.1%	-0.024 ~ -0.024	-0.04	-0.06	0.001	93.9%	-0.103 ~ -0.026
言語能力	0.03	0.04	0.006	60.5%	-0.056 ~ 0.145	0.03	0.05	0.017	33.6%	-0.121 ~ 0.215	0.01	0.01	-0.003	119.6%	0.010 ~ 0.010
非言語能力	0.06	0.08	0.009	47.8%	-0.046 ~ 0.203	0.13	0.20	0.019	29.3%	0.028 ~ 0.378	0.08	0.13	0.005	76.1%	0.036 ~ 0.217
一般知的能力総合	0.05	0.07	0.007	56.3%	-0.033 ~ 0.177	0.10	0.16	0.027	23.8%	-0.051 ~ 0.371	0.06	0.09	0.002	88.7%	0.029 ~ 0.147

2 つめのポイントは、基準変数の精度が向上した可能性である。個人の勤続年数が長くなるほど、より詳細にその職務遂行能力を観察することが可能となる。また、短期的には好環境に恵まれる等、実力以外の偶然が作用して高く評価される可能性もあるが、長期に及ぶことで偶然による要因が取り除かれ、純粋に個人の職務遂行能力が評価されるようになる。これらの結果、基準変数の精度が向上し、妥当性係数の水準を高めた可能性がある。

3 つめのポイントは、職務経験を通じて性格が妥当性係数を高める方向に変化した可能性である。Roberts, Caspi & Moffitt (2003) は、性格は職務経験を通じてその職務に求められる性格特徴が強まる方向に変化する可能性を提示している。加えて、もともとその職務に求められる性格特性を備えていた場合の方が強化される程度が大きいとしている。すなわち一般に組織の中で仕事をする上で求められる社交性、協調性、意欲といった特徴をもともと強くもっている人ほど職務経験を通じてこれらが強まる方向に変化する。その結果、妥当性係数が5年までのグループ、6~10年グループ、11~15年グループと徐々に高まった可能性が考えられる。

一般知的能力検査では性格検査と異なり、.15を超える妥当性係数は6~10年のグループのみで得られた。前述したように一般知的能力の妥当性係数の水準は職種特性の影響をうける可能性があることがこれまでの研究で指摘されている(高橋・西田, 1994)。そこで3グループの営業系、事務系、技術・開発系、その他、不明からなる職種構成を確認した。一般知的能力が特に求められる技術・開発系の職種の割合を比較してみると、5年までのグループでは18.5%、6~10年のグループでは26.8%、11~15年のグループでは16.8%となっており、6~10年のグループは他の2グループより技術・開発系職種が多いことがわかった。このことから6~10年グループの一般知的能力の妥当性係数が高くなったのは、サンプルの職種特性の影響をうけた可能性が考えられる。

5. 今後の課題

本研究は、性格検査、一般知的能力検査の妥当性について、予測的妥当性でデザインされた研究のみを使用してメタ分析的に妥当性係数の真値を推定するとともに、特に先行研究が少ない予測変数と基準変数の収集時期のひらきに注目して分

析・考察を行った。今後の課題としては以下のことが挙げられる。まず、本研究では各グループの職種構成が均質になっていないことである。また企業風土が妥当性係数の水準に影響している可能性についても検討が必要である。さらに本研究では予測変数と基準変数が収集された時期のひらきが15年までのサンプルを対象としたが、時期のひらきが妥当性係数に与える影響の一般的傾向をより明確にするためには、その先の15年以上のケースについても研究する必要がある。

《引用文献》

- Barrick, M.R. & Mount, M.K. 1991 The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.
- Hunter, J. E. & Hunter, R. F. 1984 Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- 高橋 潔、西田 直史 1994 知的能力検査に関する妥当性一般化、産業・組織心理学研究 第8巻 第1号
- 二村 英幸・今城 志保・内藤 淳 2000 管理者層を対象とした性格検査・知的能力検査の妥当性のメタ分析と一般化、経営行動科学第13巻 第3号 159-167
- 大沢 武志・芝 祐順・二村 英幸 編 2000 人事アセスメントハンドブック 金子書房
- Dreher, G.F & Bretz, R.D.Jr. 1991 Cognitive Ability and Career Attainment: Moderating Effects of Early Career Success. *Journal of Applied Psychology*, 76, 392-397.
- Hunter, J. E., & Schmidt, F. L. 1990 *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.
- 丸山 真智子・二村 英幸・今城 志保・内藤 淳 一般企業人を対象とした性格検査の妥当性のメタ分析と一般化、経営行動科学 投稿中
- Roberts, B.W., Caspi, A. & Moffitt, T.E. 2003 Work Experiences and Personality Development in Young Adulthood. *Journal of Personality and Social Psychology* 84, 582-593.