

# 客観採点式インバスケットテストの開発と妥当性検証

○堀博美 今泉緑

(日本エス・エイチ・エル(株))

Development and validation of the objectively scorable in-basket test

Hiromi Horii Midori Imaizumi

(SHL-Japan Ltd.)

## 背景

インバスケット演習とは、日常業務で使用されるような書類に準じた模擬書類が作成され、未決裁箱に入ったそれらの書類に対して一定時間内に検討を加え、決断を下していく模擬演習である。

1953年アメリカ空軍大学の教育効果測定を目的に開発され、すぐに産業界に応用された。日本に紹介されたのは1970年代である(楨田, 1988)。高度経済成長が一息つき、これからは経験とバイタリティの鬼軍曹的なリーダーよりも頭を使う管理者を求めようとする気運に合致し、一部の企業で管理職登用などの場面で活用された。初期のものとして慶応ビジネススクールにおける関本ら(1971)の研究によるKBS式がよく知られている。

当初は主に管理能力の評価・開発を目的とし、昇進昇格場面でアセスメントセンター(以下AC)などでよく利用されてきたインバスケット演習であるが、近年、この演習で測定されるような能力を管理職層よりもかなり前の段階で明らかにしたいとのニーズが高まっている(マネジメントサービスセンター, 1994)。すなわち、業務の高度化・IT化を背景に、特にホワイトカラー・ビジネスマンにとって、膨大な情報をその根底にあるつながりからテーマ別に整理し、解決すべき案件を見分け、優先順位をつけながら合理的な判断を下していく能力は非常に重要であり、たとえば採用

段階においても可能ならば評価したい、というニーズである。

このニーズに対し、インバスケット演習の最大のデメリットは、採点評価が難しくかつ時間がかかるという点になる。演習の実施や評価の方法はさまざまに工夫されているが、基本的に受検者は意思決定のプロセスや内容を記述式で回答する。評価者は記述された回答をその内容やスタイルの観点であらかじめ用意された評価基準書に基づいて採点する。ただ、評価基準書はあるもののあくまでガイドラインであり、受検者が書いた内容がそれに合致するかどうかなど評価者のスキルや主観的判断が入る余地が大きい。

この欠点を補うため、我々は選択肢形式で回答させ、コンピュータによって客観採点できるインバスケットテストを2001年に開発した。本発表の目的は、この開発のプロセスを報告し、現在までに収集されたデータによって妥当性を検証することである。

## インバスケットテストの開発

【時期】2000年11月～2001年5月

【手順】開発プロセスは図1のとおりである。

【概要】

① 筆者らが所属する機関は、様々な企業に人材アセスメントサービスを提供している。1998年頃からAC形式の研修サービスを手がけ、その中で記述式のインバスケット演習を実

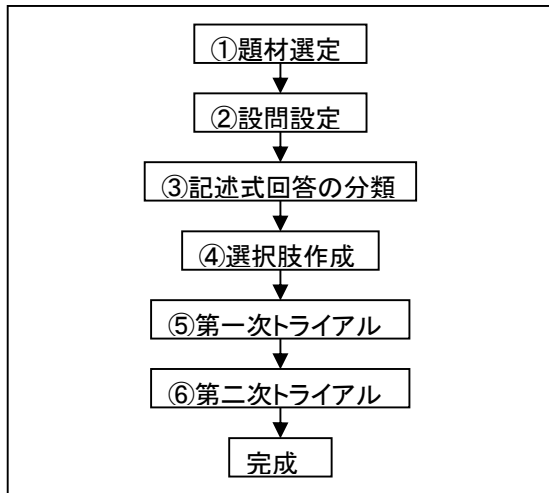


図1 開発プロセス

施しているが、回答データがある程度蓄積されたもので、受検者が担わされる役割のレベルが比較的低く、職務経験のほとんどない受検者でも回答できるような題材を選定した。

② 記述式で実施していた時も、回答のさせ方は比較的構造化されていた。すなわち、20～30の資料を与え、(ア)複数の資料を関係づけながら着手すべき案件を識別して優先順位と行動計画を策定させる；(イ)題材に含まれる重要な問題を提示し、資料をもとにその解決策を根拠と共に詳しく記述させる；(ウ)クレームに対する返信など文章を書かせる；(エ)予算や人員計画など数値の含まれる課題を解決させる；の4つの課題があり、それぞれに制限時間を設けて回答させる形式であった。これらのうち(ウ)は選択式に向かないと判断してはらずし、残りの課題から設問を設定した。

③ 蓄積された記述式での回答データを収集し、各課題に対して実際に受検者がどういう回答をしたかを分類・整理した。

④ 能力検査問題作成の専門家で職務経験豊富な者3名がそれぞれ、③で整理したデータをもとに②の設問に対する選択肢原案を作成。弊社研究開発担当者2名がそれらの原案を検討し、トライアル用の選択肢をまとめた。さらに、記述式での評価基準書を念頭に妥当解

を設定した。

⑤ 弊社で主催する面接コースの被面接者(就職活動中の大学生)24名に第一次トライアルを実施。項目分析として、当該項目とテスト総合得点の相関係数を算出した。係数が0.2以下の項目について再度検討を加え、加筆・修正した。

⑥ 弊社採用説明会に参加した大学生168名に第二次トライアルを実施。再度項目分析を実施したところ、⑤で修正した項目について0.2以上の値が確保された。標準化に十分な得点分布が得られたことから、素点得点を1～10のステン得点に変換した。

### インバケットテストの妥当性検証

#### 【目的】

客観採点式インバケットテストの妥当性をさまざまな角度から検証すること。

#### 分析① 能力検査との関係

【実施時期】2003年

【サンプル】情報A社採用応募者7544名。

【手順】3種類の能力検査(言語理解、計数理解、構造理解；日本エス・エイチ・エル株式会社)をインバケットテストと同時に実施した。

【結果と考察】インバケットテスト得点と能力検査3尺度およびその総合得点の尺度間相関は表1のとおりであった。

サンプル数が多いため値は全て1%水準で有意であるが、インバケットテストと能力検査の相関係数は0.2前後である。

関本(1979)は記述式インバケット演習と管理職適性検査MATの相関を分析し、MAT知的能力との相関係数が0.34～0.36であったことを報告している。本研究の値はそれよりもやや低めだが、その理由がインバケットの記述式と選択式という回答形式によるものかどうかは、演習題材や能力検査の測定内容も異なるため、さらに検証が必要であろう。

いずれにしても能力検査との相関はさほど高くない。インバスケは実務処理能力を測定することを意図しており、全般的な知的能力とはやや異なる側面を測定していると考えられる。

表1 能力検査との相関 ( $n = 7544$ )

	インバスケ総合	言語	計数	構造	
インバスケ	1				
総合	0.22	1			
言語	0.15	0.62	1		
計数	0.16	0.77	0.17	1	
構造	0.14	0.70	0.15	0.39	1

(値は全て $p \leq 0.01$ で有意)

### 分析② パーソナリティ検査との関係

【実施時期】と【サンプル】は分析①と同じ。

【手順】パーソナリティ検査 OPQ (日本エス・エイチ・エル株式会社) をインバスケテストと同時に実施した。OPQ は 30 尺度を測定するイプサティブ形式の検査である。

【結果と考察】インバスケテスト得点とパーソナリティ 30 尺度の相関のうち、1%水準で有意な 11 尺度を表 2 にまとめた。サンプル数が多いため有意ではあったが、値は極めて小さい。

表2 パーソナリティ検査との相関 ( $n = 7544$ )

	インバスケ
指導力	0.05
謙虚さ	-0.04
面倒み	0.03
具体的事物	-0.05
美的価値	-0.04
概念性	0.03
創造的	-0.02
タフ	-0.02
批判的	0.03
行動力	-0.03
決断力	0.04

(値は全て $p \leq 0.01$ で有意)

外島 (2001) は AC ディメンションの「問題把握の因子」(分析力、理解力、計画力、判断力など) を目的変数、TPI 尺度因子得点を予測変数とした重回帰分析を行ったが、有意な重

相関係数が得られなかった。また、関本 (1979) も記述式インバスケ演習と MAT 性格尺度との間に低い相関しか見られないと報告している (性格的強靱性と 0.06~0.10; 決断性 0.14; 支配性 0.13~0.14; 社交性 0.19~0.22)。本研究でもインバスケテスト結果とパーソナリティの間にはほとんど関係がないという結果になった。

### 分析③ 職種との関係

【時期】2004 年~2008 年

【サンプル】2 社 1420 名。初級・中級管理職候補者。平均年齢 39.2 才 (範囲 32 才~54 才)。

【手順】受検者を営業・スタッフ・技術の 3 つの職群に分け、それぞれのインバスケテストの平均点を算出した。

【結果と考察】結果は表 5 と図 3 のとおりである。インバスケテストの平均値は、スタッフ職 > 技術職 > 営業職であった (営業とスタッフの間に有意差 ( $p \leq 0.01$ ))。スタッフ職では書類を扱うデスクワークが多く、業務で求められる要件から納得できる。

表3 職種別傾向 ( $n = 1420$ )

職種	営業	スタッフ	技術
$n$	755	319	346
平均値	6.04	6.52	6.32
標準偏差	1.87	1.84	1.78

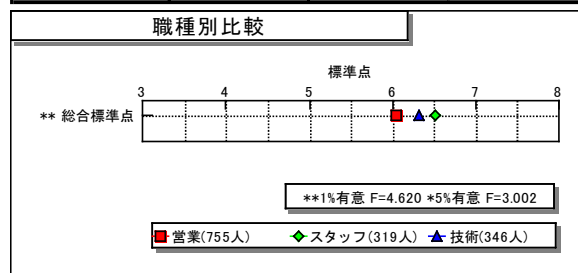


図2 職種別傾向 ( $n = 1420$ )

### 分析④ ACにおける行動評価結果との関係

【実施時期】2005 年~2007 年

【サンプル】5 社 128 名。初級管理職及びその候補者で業種や年齢は様々。年齢は主に 30 代~40 代前半。

【手順】サンプルは弊社が実施したACの参加者である。ACではグループ討議など4~5の模擬演習を実施し、6つのディメンションについて複数のアセッサーがその行動を5段階で評価している。

【結果と考察】インバスケットテスト得点とAC行動評価6ディメンションの相関を表4に示す。

表4 AC行動評価との相関 ( $n = 128$ )

	インバスケット
リーダーシップ	0.02
チームワーク	0.15
対人積極性	0.12
分析力	0.41 **
創造力	0.24 **
手順化能力	0.10

(\*\*  $p \leq 0.01$ で有意)

インバスケットテストの得点は、AC行動評価の分析力、創造力と有意な相関が見られた。最も相関の高かった分析力の評価段階別にインバスケットテストの平均値を算出し、比較したものが表5と図3である。

表5 AC(分析力)評価との関係 ( $n = 128$ )

AC評価	1	2	3	4	5
$n$	17	31	39	21	20
平均値	4.82	4.87	5.72	6.38	6.90
標準偏差	2.16	1.69	1.54	1.56	1.86

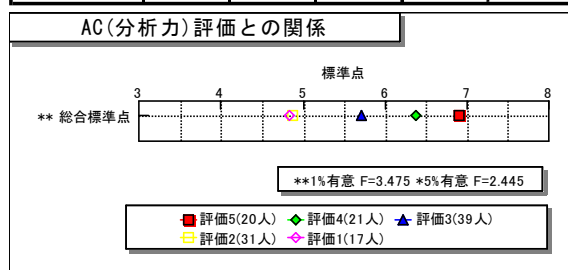


図3 AC(分析力)評価との関係 ( $n = 128$ )

関本(1979)は記述式インバスケット演習とMATと人事評価の関係分析から、人事評価の上位群の識別にインバスケット、下位群の識別に心理テストが有効であると結論付けている。(人事評価と行動評価という点で評価内容は異なるが、)インバスケットテストは段

階1と2という評価の低いところは識別できていないが、段階2以上の評価については識別できており、記述式演習と心理テストのほぼ中間に位置しているとも考えられる。

### 今後に向けて

企業組織で実際に活用されるテストの妥当性で最も重要なものは人事評価との関係であろう。本発表では情報管理の点でデータを掲載できなかったが、実際には数社で分析を行い、人事評価と関係していることが確認されている。さらに多くの企業、幅広い業種でデータを蓄積していきたい。

幹部社員の早期選抜は重要なテーマであるが、早期であればあるほど母集団は大きくなる。ある程度の規模の受検者に実施・採点するコストを考えると、実用場面においてこの客観採点式インバスケットテストは非常に有望である。さらに、現実の仕事場面ではPC上で業務を進めることが主流になっていることも考えると、PC上で実施させるインバスケットテストの開発が考えられる。また、Eラーニングへの活用も期待される。

### 参考文献

- 榎田仁(1988) 管理能力開発のためのインバスケット・ゲーム 金子書房
- マネジメントサービスセンター編(1994) 人材の発見と選抜—ヒューマン・アセスメントはここまで広がった— ダイヤモンド社
- 関本昌秀・平尾洋一・畠山泰雄・望月良三・今野勝利(1971) 管理能力の発見とインバスケット・テストI—インバスケット・テストの意義と構成について— 慶応ビジネスフォーラム 10 1-33
- 関本昌秀(1979) 管理能力の発見とインバスケット・テストIII—MATとの比較における妥当性の検証— 慶応経営論集, 1(2), 1-20.
- 外島裕(2001) アセスメントセンターのディメンション因子構造とパーソナリティ特性 日本性格心理学会大会発表論文集, 10, 58-59.