

知的能力検査の問題項目の 作成と分析事例

RECRUIT
MANAGEMENT SOLUTIONS

藤田 彩子
株式会社リクルートマネジメントソリューションズ
組織行動研究所 研究員

はじめに ~ テスト作成の難しさとは ~

例えば小説を書くことを例に考えると・・・

- ・何を問題にしているのか
- ・どう問題にすればいいのか
- ・これでいいのか
- ・これでよかったのか

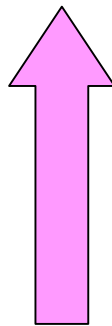
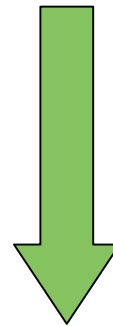
ネタさがし

手法・テクニック

推敲

読者からの評価・受賞

ノウハウになりやすさ



オリジナリティの発揮

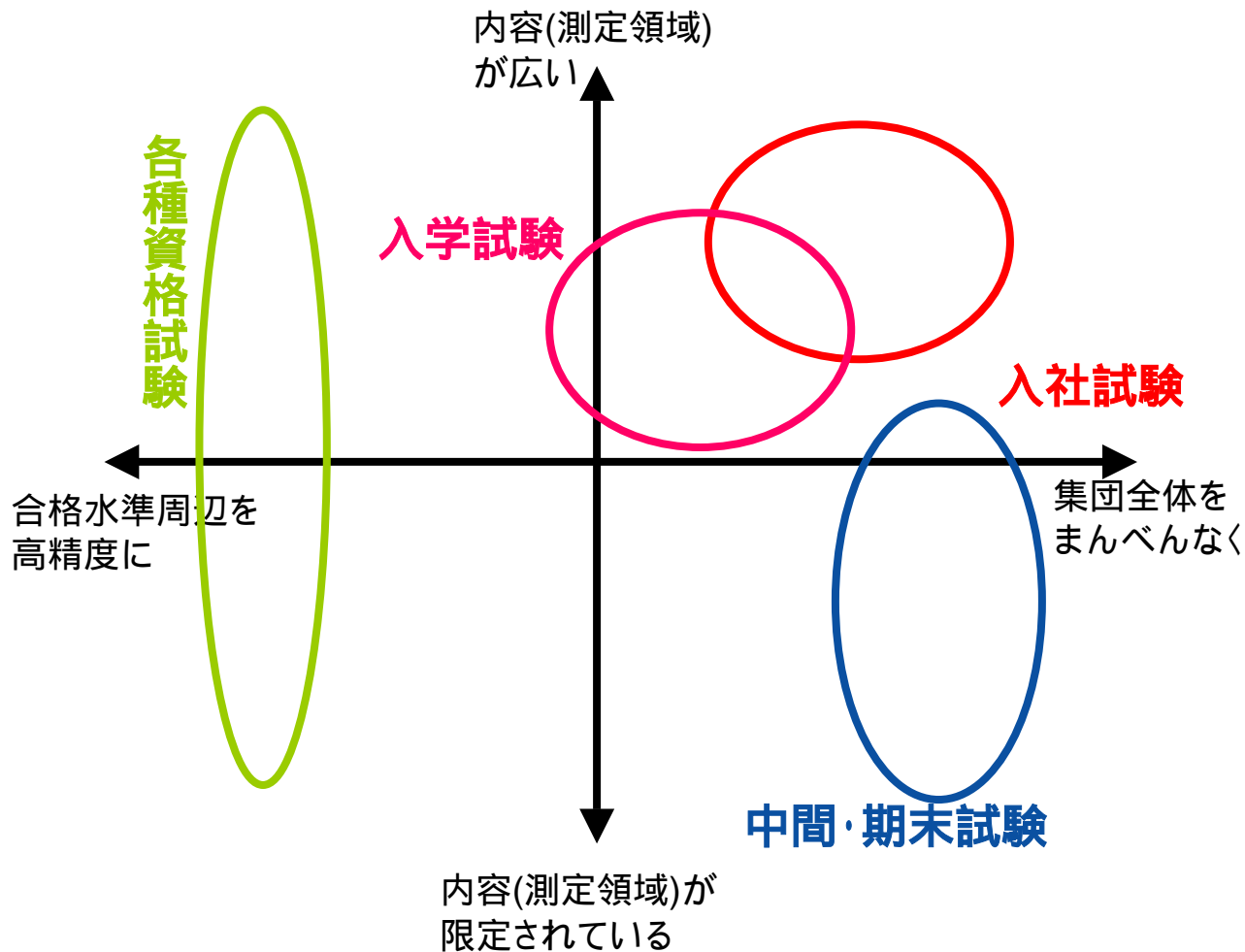
問題作成ってある意味、「作品を創る」ようなものですよね。

その他独特の問題として、

問題項目の管理、分析の仕方・・・

はじめに ~つくりたいのはどんなテスト?~

世の中にはどんなテストがあるでしょうか?
それぞれのテストの目的・用途は?



縦軸:測定領域の広さの観点
漠然とした広い領域(ex.言語能力)を測定したいのか,限定された領域(ex.漢字の書き取り)を測定したいのか
問題の選び方・結果の見方に関わる

横軸:得られる情報の精度の観点
集団における位置づけが目的なのか,ある決まった基準を満たしているかどうかを厳密に見たいのか
テスト情報量の設計に関わる

まず,テストの目的をはっきりさせることが重要

1. 小社が提供している「一般知的能力検査」の概要
2. 問題項目の作成からテストの完成までの流れ
3. 問題作成～編集～データ分析の各ステップにおける留意点
4. 課題・まとめ

リクルートマネジメントソリューションズ概要

■ 会社概要

| | |
|-------|---------------------|
| 創 業 | 1963年 |
| 設 立 | 1989年 |
| 資 本 金 | 1億10百万円 |
| 売 上 高 | 133億72百万円（2004年度実績） |
| 従業員数 | 384名（2005年4月現在） |

■ 事業概要

- 人事アセスメント事業
- 経営人事コンサルティング事業
- トレーニング事業

■ 総合検査SPI2(Synthetic Personality Inventory)

- おもに採用選考場面で人物の総合的理解をサポートするツールとして利用
- 「性格的側面」と「能力的側面」で構成される適性検査

■ CBTサービス展開

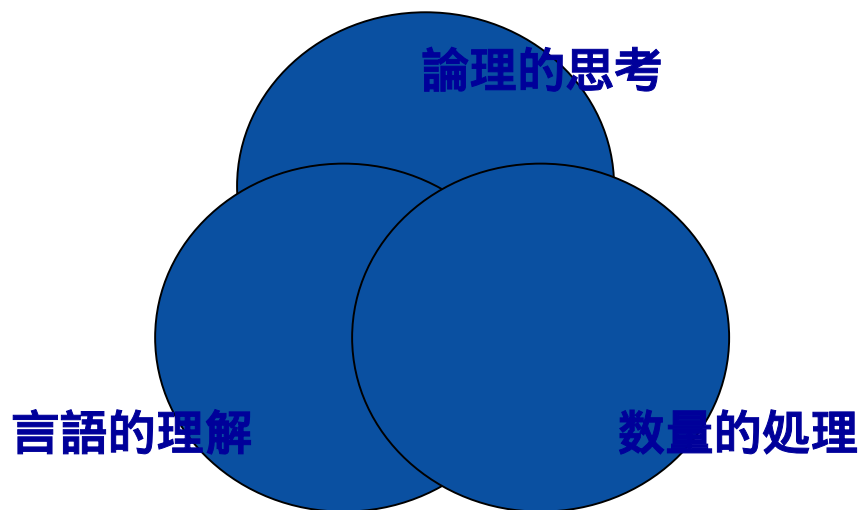
インフローテストティングサービス（2004年1月～）

- SPI2を適応型テストとして提供
- そのほかWebでもSPI2を提供

職種の違いを超えて共通に要求される能力を測定する

具体的には、

- ・新しい環境や課題に対して合理的に思考し、目的を定めそれに沿って行動し、効果的・効率的に事態を処理していく能力
- ・複雑なことを整理したり判断したりする能力
- ・新しいことを学んだりもっている知識を応用したりする能力



一般知的能力は、
「言語的理解」
「論理的思考」
「数量的処理」
の3因子から成り立っている

“一般知的能力検査”の紹介(技術面)

■ 客観式

■ 規準集団をもとに得点が標準化されている

背景となっている集団の得点分布をもとに個人を位置づける

50点なら規準集団の上位50%の能力を持つ,
60点なら上位約15%の能力を持つ, など



- ・得点に意味を持たせることができる
- ・平行版がある場合, 結果を同一の基準で比較することができる

■ 複数の平行版があり, 毎年少しずつ改訂している

問題も平行問題を作成

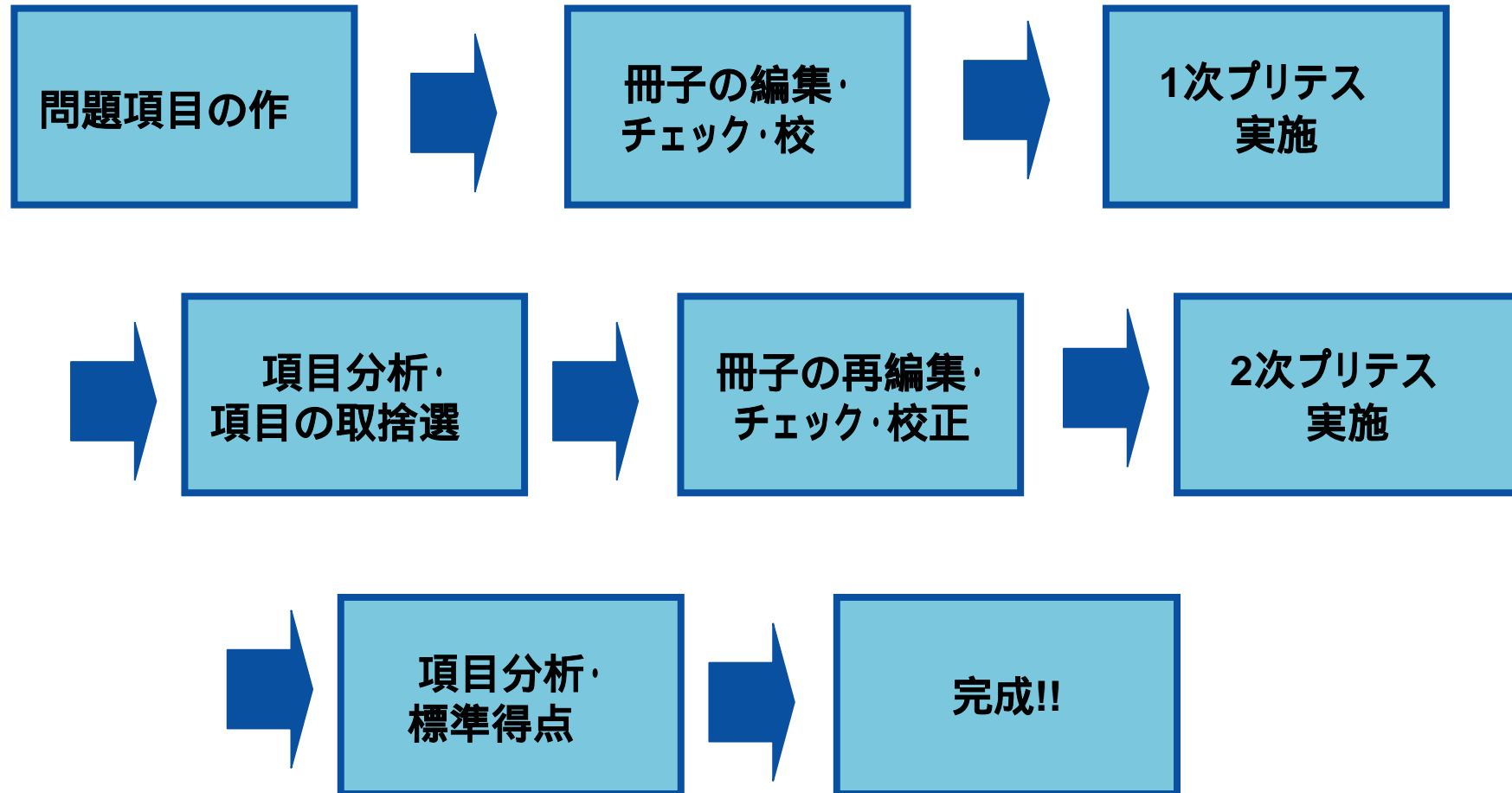
■ 媒体は, P&P(複数の平行版), CBT(適応型)の2種類

P&Pは各版で素点-標準得点の換算表をもつ

CBTはIRTを用いて採点

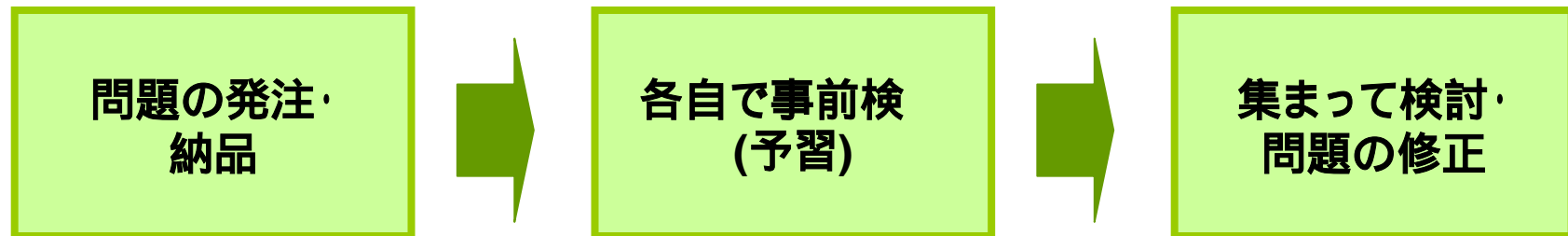
1. 小社が提供している「一般知的能力検査」の概要
- 2. 問題項目の作成からテストの完成までの流れ**
3. 問題作成～編集～データ分析の各ステップにおける留意点
4. 課題・まとめ

テスト作成の大きな流れ



問題項目の作成

言語検査・非言語検査それぞれ10人弱の作成チームがあり、定期的に項目作成検討会を行っている。



目的:作成された問題案について,主旨・観点・表現が適切かどうか複数名で検討し,必要な修正を施す

問題作成において,検討会は何よりも重要!!

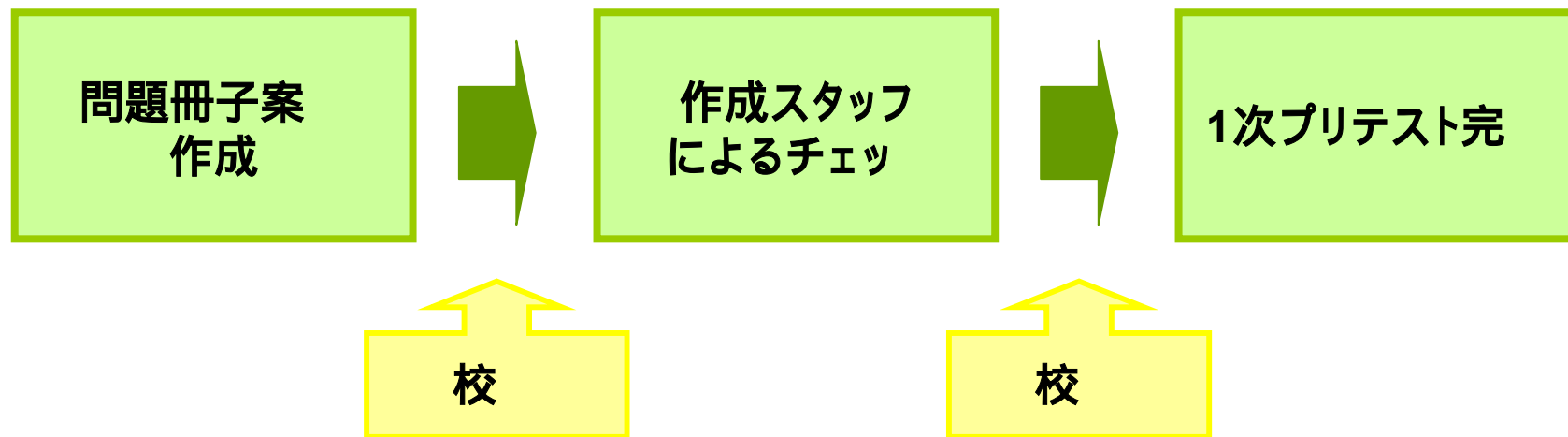
どんなに問題作成のベテランになっても,自作の問題を様々な視点から,客観的に見ることは難しい。

様々な視点を持った複数の人たちの意見を取り入れて問題を完成させることが重要。

問題作成のスキルが伝達される場でもある。

1次プリテストの作成・実施

これまで作成した問題を組み合わせる冊子案とし、社員・作成スタッフによる校正・チェックを経て1次プリテストが完成する



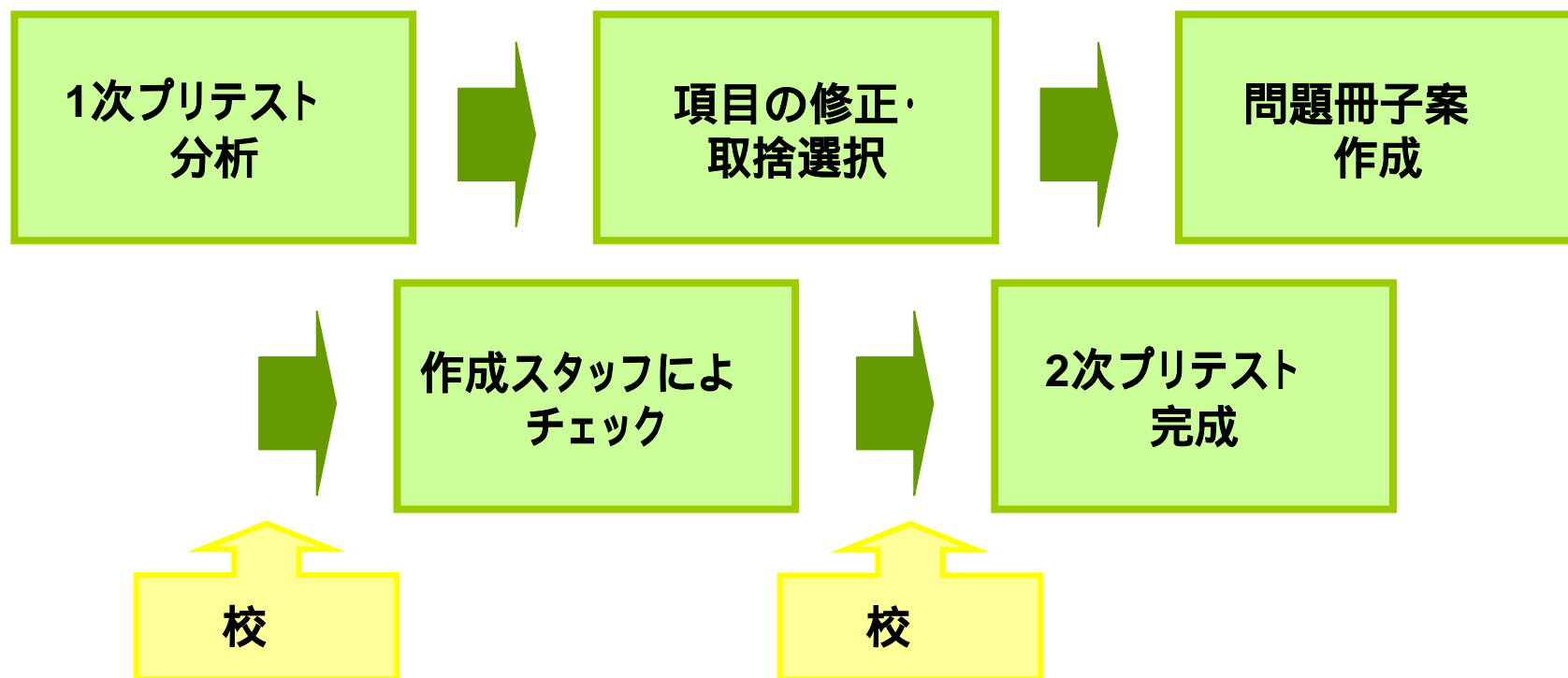
1次プリテスト実施の目的

目的:問題のよしあし, 選択肢のよしあしなどをチェックする

検討は済んだが, まだ難易度・識別力等の素性の分からない問題についてデータ収集をして, 識別力の悪い問題や有効でない誤答選択肢を捨てる

あとから問題や選択肢を取捨選択できるように, 問題・択肢とも少し多めに用意する

1次プリテストの結果を踏まえて項目を取捨選択・修正し、新たに2次プリテストとして編集しなおす



2次プリテスト実施の目的

**目的:商品としてすでに提供されているテストと等化し,標準
得点化するために必要なデータを集める**

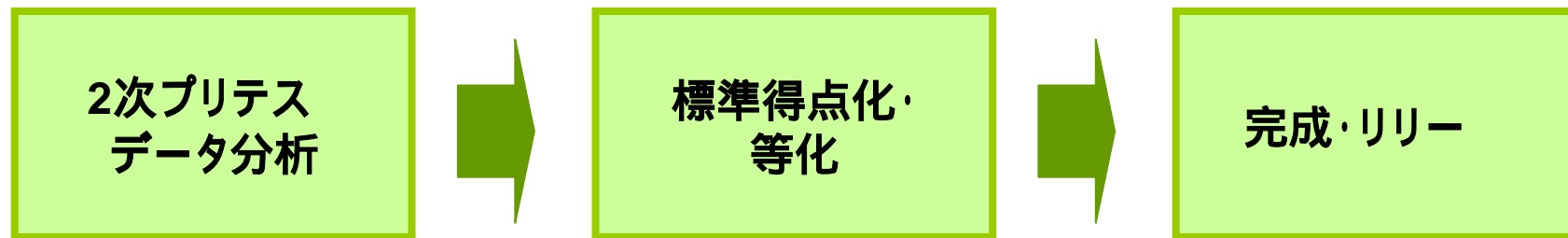
**1次プリテストでOKとなった問題を集め,2次プリテスト
として編集しなおす**

**2次プリテスト作成が完了した段階でテストとして内容
確定し,その後は原則的に問題の内容の変更や問題
の差し替えなどは行わない**

**使用する問題項目の難易度・共通項目の選定などに注意を
払って編集する**

商品として完成・リリース

共通項目の情報を元に既存の版と等化し、
テストとして完成する



その後も毎年等化状況を確認して、必要に応じて素点-標準得点の換算表の微調整を行い、どの版を受検しても同じ基準での標準得点が報告できるようにしている

1. 小社が提供している「一般知的能力検査」の概要
2. 問題項目の作成からテストの完成までの流れ
- 3. 問題作成～編集～データ分析の各ステップにおける留意点**
4. 課題・まとめ

問題作成・検討会のポイント 1

■ 作成基準のすりあわせ

問題作成ガイドラインに問題形式・作成ノウハウをまとめ、皆が同じ基準で問題作成ができるようにする

問題作成ガイドラインとは:

- ・問題作成・検討を経てテストとして完成してゆくまでの流れを解説
- ・問題の各形式について、測定領域の定義(何を測定する問題か)、例題、作成における注意点や作成のコツなどを明記

新人スタッフの導入に役立つだけでなく、中堅・ベテランスタッフも、作成のたびによりどころとしている

■ 作成スタッフと運営スタッフとの役割分担

運営スタッフは議論が偏ったり迷走したりしないようコントロールし、テスト開発の主体として随所随所で議論に決着をつけていく

■ 活発な検討が行われるような雰囲気作り

特定のメンバーだけに発言が偏らず皆が遠慮なく発言できるような和やかかつ緊張感のある雰囲気を作る

■ 検討は、1回で終わりにしない

- ・1回目の検討では修正がたくさん出て問題が大きく変わってしまうことも多い
 - ・同じ問題に長時間向き合っていると、考えすぎ・検討しすぎた結果、かえってよく分からない問題になってしまうこともある
- 修正部分をきれいに清書したものを、冷却期間をおいてから冷静にもう一度検討してみる

■ ある程度完成した問題を、いくつも並べて相対的に見てみる

難易度調整やそぐわない問題を修正するなどして完成させる

よい問題を作るコツ

原

- ・問題の意図をはっきりさせる(問題とすべきポイントは何か)
- ・ポイントを的確に問うのに適した形式・体裁を考える
- ・表現は分かりやすく簡潔か、あいまいなところはないか
- ・一つの問題・一つの選択肢のポイントは一つに絞る
- ・取りようによっては複数の解釈がある表現は避ける

できるだけ気をつけたいこと

- ・選択肢の「レベル」をそろえる
- ・選択肢の長さ・言い回しなどをそろえる
- ・「必ず～」「いつでも～」といった表現を使うことで誤答とする等の安易なやりかたは避ける
- ・(問題によっては)正答とすべき内容を正答とし、受検者の納得感を高める

1次プリテスト作成のポイント

■ 難易度や内容のバランスを見ながらテストとして編集する

- ・ただし，難易度については感覚でしか分からないのであまり凝ってもしかたない
- ・最初は比較的易しめの問題を用意し，受検者がスムーズにテストに入れるようにする

■ 後々の分析の時に役立つように，特性の分かっている項目も少し混ぜておく。

データ分析の結果が項目によるものなのか，受検者の集団の影響によるものなのか考える際の参考になる

(ちょっと横道) 校正に関して

■ 表記の統一

- ・ 表記統一のためのマニュアルを作成し、複数の校正者が同じ基準で校正できるようにする。
(例外: 文章読解問題など, 原文を尊重するべきと判断される場合)

■ 見やすいレイアウトを心がける

- ・ 選択肢の順番はできるだけ見やすく, かつ恣意的にならないよう, 形式的な並べ方をする

1次プリテストの分析

■ 項目分析・選択肢分析

項目のよしあし，難易度，有効/無効な選択肢を判断する

項目の選定(データの悪い項目はボツに)

有効でない，または「有害な」選択肢の削除

気をまわしてあれこれ邪推す

- A みなす
- B ねたむ
- C かんぐる**
- D さつする
- E あやしむ

N=6604

| (%) | 選択肢A | 選択肢B | 選択肢C | 選択肢D | 選択肢E | NA |
|-----|------|------|------|------|------|-----|
| 上位群 | 0.1 | 0.2 | 95.6 | 1.1 | 3.0 | 0.2 |
| 中位群 | 0.2 | 1.3 | 87.5 | 4.3 | 6.3 | 0.4 |
| 下位群 | 1.0 | 5.4 | 69.8 | 11.5 | 11.3 | 1.0 |

正答は，上位群が多く選択し，下位群の選択率と大きな差がある

先例に外れるこ

- A 論外
- B 無比
- C 破格**
- D 例外
- E 破天荒
- F 未曾有

N=2508

| (%) | 選択肢A | 選択肢B | 選択肢C | 選択肢D | 選択肢E | 選択肢F | NA |
|-----|------|------|------|------|------|------|-----|
| 上位群 | 0.4 | 2.3 | 8.1 | 38.0 | 11.9 | 38.3 | 0.9 |
| 中位群 | 4.2 | 3.7 | 4.9 | 55.2 | 9.3 | 21.7 | 1.0 |
| 下位群 | 14.3 | 3.5 | 5.5 | 58.0 | 7.1 | 10.8 | 0.9 |

正しくない選択肢を上位群が多く選択し，誤答にもかかわらず正答と同様のパターンを示している

■ 問題作成のベテランスタッフのノウハウを どうやって経験の浅いスタッフに伝授していくか

- ・ **ガイドラインなどのマニュアルの整備**
「ガイドラインを作成する」ことによる、ノウハウの言語化や作成にあたっての個々人の工夫の共有の効果は大きい
- ・ **問題検討が一通り済んだあとで、今回の作成問題
に関して感想・意見を交換**
できれば記録をとって、次回似たような問題を作成するときの参考にする

結果のフィードバックの重要性

- テストとして編集し，実施したら，分析結果を作成スタッフにフィードバックすることが重要
 - ・ 自分たちが作成した問題の結果を共有することで作成に対する意欲を高める
 - ・ 分析結果は，難易度のレベル感が分かる・有効だった誤答・有害だった誤答・無効だった誤答の傾向が分かる，など今後の問題作成に役立つ情報の宝庫である

- 作成スタッフの確保・育成
- 効率と品質との折り合い
- 安定性を損なわずに陳腐化を防ぐ

- 検討会の重要性
- 作成スタッフの積極的な姿勢を引き出す
- 作成ノウハウの伝授・言語化
- 結果のFB
- 効率化と品質
- マニュアル化とマインドの伝達
- データによる判断と人による判断