

全国的な学力調査のCBT化検討WG(第9回)
令和3年4月23日 10:00~12:00
Web会議

IRTを採用した調査に必要な 実施体制等について

(株) ハピラル・テストソリューションズ
代表取締役社長 別府正彦



本日のお話

- 私どもは、お客様が保有しておられる「テストに出題した問題」と「実施したテストのデータ（受験者の正誤データ）」を用いて、IRTを採用するテストの開発をお手伝いしています。自社で独自にIRTを採用したテストを開発しているわけではありません。
- 「必要な体制」などについて一例をお話ししますが、実情に応じて、それ以外にも様々な方法が考えられ、答えは決して一つではありません。また、IRTに関するお話も致しますが、時間が限られておりますので、簡便な説明になろうかと思えます。これらの点を、ご理解いただければ幸いです。
- 本日は、拙著「IRT入門」の内容を中心に、普段の業務を通じて経験したこともふまえて、次の3点についてお話しいたします。

1. IRT：導入のメリットとよく耳にする誤解
2. 問題セットの作成に必要な体制
3. 日々の学習での活用例

～どうぞよろしくお願いたします～





IRT導入のメリット

出題する問題が異なっても、児童・生徒の得点を比較できる。

- 複数種類の問題セットの難度を、ほぼ同じ程度に揃えることが可能
- 複数種類の問題セットを使い、複数日程で調査を実施することが可能
- 問題セットの種類を増やし、教育課程をより広くカバーすることが可能
- 学力の変化を時系列的に追跡することが可能

← 共通尺度

← 等化

※以下、IRTを導入した（する）調査のことを、「IRT調査」と呼びます。



IRTについて耳にする誤解～1

1) IRTを導入すると大幅にコスト高となる！

- IRTを導入すると、多くの場合、次の業務が増えます。

IRTによる分析・等化
項目プールの構築
項目プールの管理・運用

- 業務が増えますので、IRT調査は、時間、人手などのコストが余計にかかる可能性があります。以前、「IRTを導入するには、事前に数万題も問題を用意し、項目プールと呼ばれるデータベースが必要だ!」、という意見を耳にしたことがあります。しかし、どの程度の規模の項目プールを用意するかは、ケースバイケースです。調査の実施目的に応じて計画を策定し、計画に沿って「最適規模の項目プール」を構築し、過剰な経費をかけないことが必要です。
- 無駄なコストを省いてIRT導入のメリットを十分に具現化するには、事前に「各種計画を立案」し、計画を実現するための「体制を整備」することが重要だといえます。



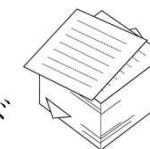
IRTについて耳にする誤解～2

2) IRTを導入する調査に記述式の問題は出題できない。



- 記述式問題のように部分点が生じる場合、段階反応モデル等で対応できます。ただし、採点基準に基づいて、正確に採点されていることが条件です。

3) IRTを導入するには数万題の問題を用意しなくてはならない。



- IRTを採用する場合、事前に問題を用意しておくことはありますが、前ページのスライドにあるように、数万題もの膨大な数の問題が必要かどうかはケースバイケースです。

4) IRTを導入する場合、問題は全て非公開にしなくてはならない。🔒

- IRTを使うからといって、必ず全ての問題を非公開にしなくてはならない、というわけではないと思います。もちろん、問題漏洩によって調査に影響が生じるようであれば、公開した問題を再出題するルールを整備したり、効率的に等化済み問題を増やす仕組みをつくったり、学力調査と競争試験は違うという考え方を周知させたりする等の対応が必要です。



IRT調査の業務の流れ

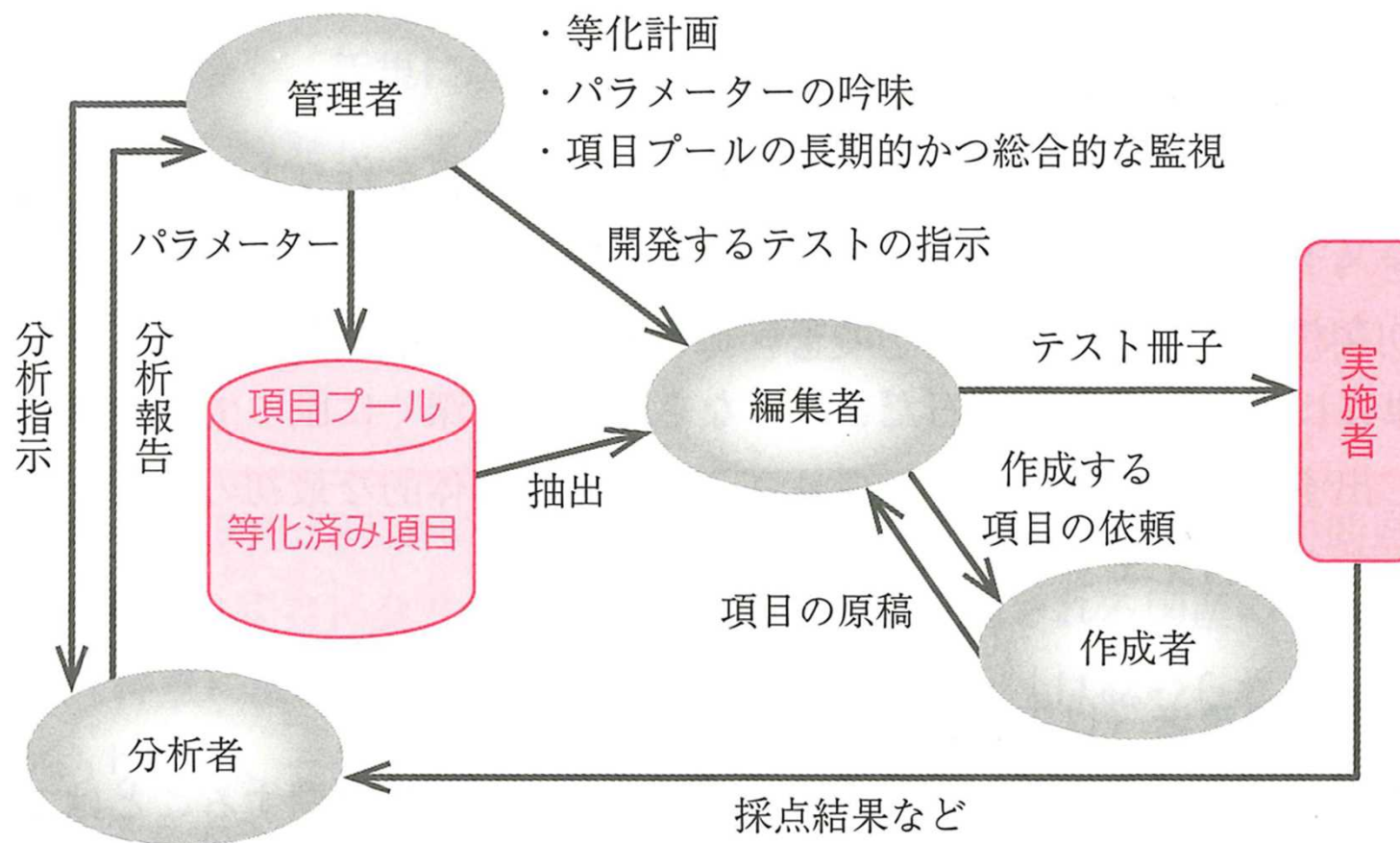
- 1) 調査を実施する目的の明確化
 - 2) 目的を達成するための計画立案*と体制整備
 - 3) 問題セットの作成
 - 4) 調査実施
 - 5) 分析・等化
 - 6) 結果の活用
 - 7) 検証
- ⇒ 3)～5)を「調査の実務」と呼びます

*：調査の出題計画、等化計画、実施計画、分析計画、活用計画、収支計画など

IRTを導入する場合、3) 問題セットの作成 は、4) 調査実施 や、5) 調査の分析・等化 と深く関係しますので、ここでは、3)～5)をまとめて、「調査の実務」とし、以下、必要な役割や体制について取りあげます。



調査・テストに必要な役割



各役割に求められること

役割	技術			知見				資質	
	分析	編集	作問	測定理論	教科知識	教育情勢	他	共通	個別
管理者	○	△	△	○	○	◎	教育に関する総合的な知見	コミュニケーション力 適度な柔軟性	・判断力 ・計画性
編集者	○	◎	△	△	○	○			・バランス感覚 ・調整力
作成者	—	—	◎	△	◎	○			・協調性 ・教科内容への考察力
分析者	◎	—	—	◎	△	○			・探究心 ・現実感覚

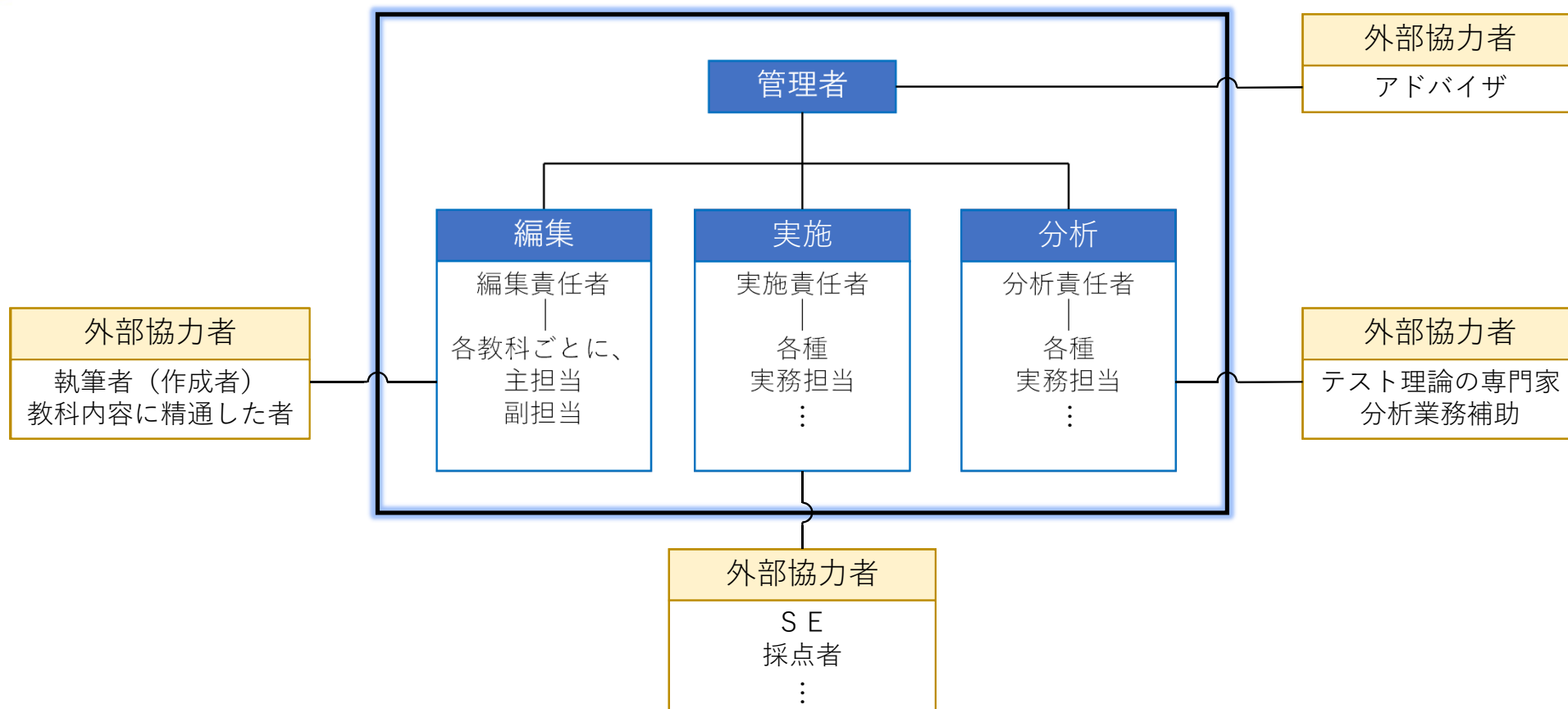
補足：記号は、技術や知見のそれぞれにおいて、各役割にどの程度求められるかの目安。

- ◎ 強く求められる
- 求められる
- △ 最低限求められる
- なくても良い





調査の実務に必要な体制





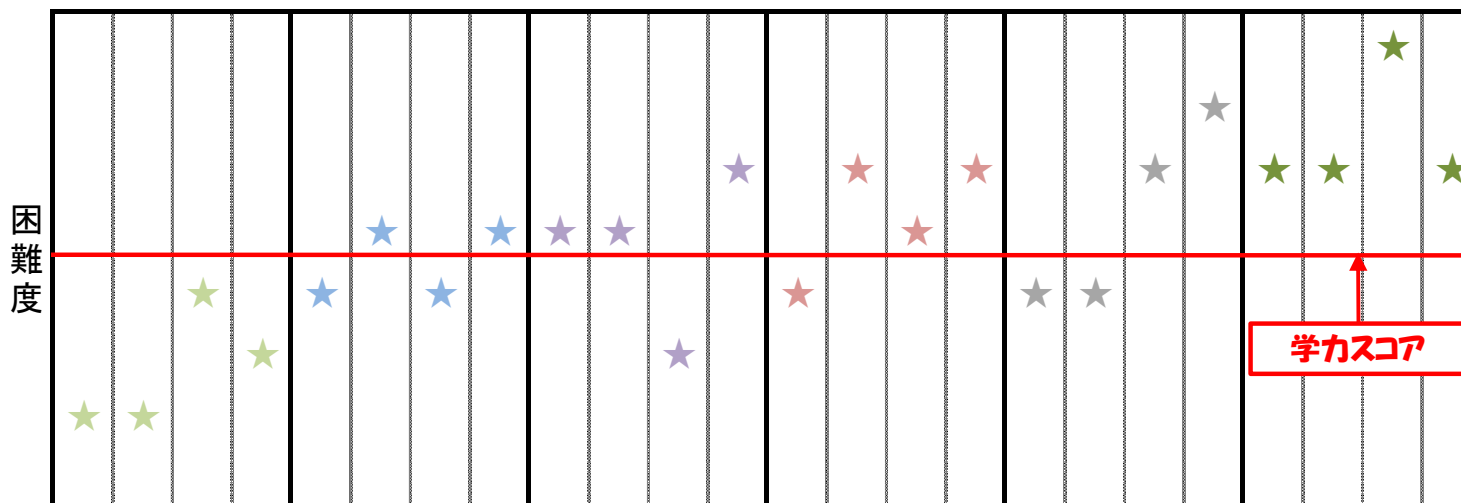
問題漏洩について

- IRT調査において、過去に出題した問題を再利用する場合、調査実施前に問題漏洩を完璧に防ぐことは困難です。
- 問題漏洩の影響を小さくする対策として、次の①～③が考えられます。
 - ①調査に出題済みで、調査とは別の場で等化した問題を大量に保有しておく
 - ②調査に未出題で、調査とは別の場で等化した問題を相応数保有しておく
 - ③調査に未出題かつ未等化の問題を、調査で等化する

補足：上記①②の場合、大量の等化済み問題の中から、どの問題が調査に出題されるかわからないので、影響を抑えることができます。
- 最も実現してほしいのは、全国学力・学習状況調査は文字通り「学力調査」であって、「競争試験」ではないという考え方が一般的になり、過度な事前対策は意味がないという考え方があまねく浸透すること、です。そうした状況になれば、上記①②は適切な数以上に大量の問題を保有する必要はなくなり、あえて③に取り組むこともありません。



日々の学習での活用例



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
○	○	○	×	○	×	○	×	○	○	×	×	○	×	○	×	○	○	○	○	○	×	×	○

分野1	分野2	分野3	分野4	分野5	分野6
-----	-----	-----	-----	-----	-----

×	現在の学力で解けるはずなのに間違えた問題 ⇒この問題から復習
○	現在の学力以上の力が発揮できた問題



※問題の困難度に関する情報と児童生徒の回答結果から得られる学力スコアを照らして、より効果的な復習や宿題を項目プールから提示することが可能になる



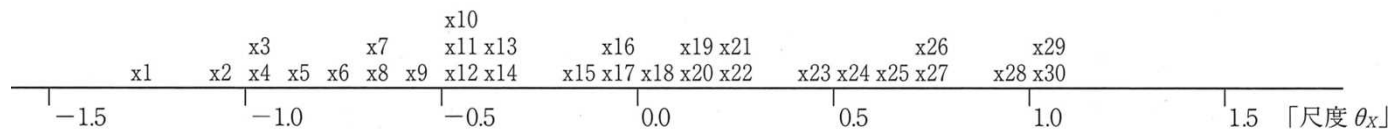
(参考) 等化とは？



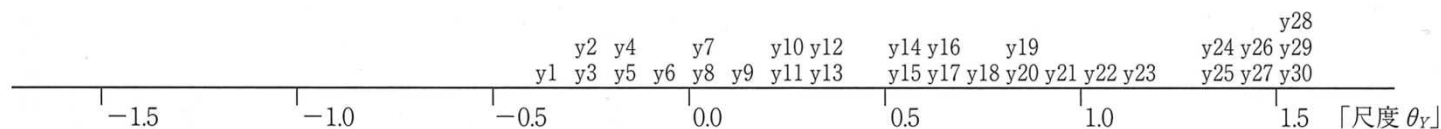
- p3のスライドにあるIRTのメリットを実現するには、「等化」と呼ばれる手順が必要です。
- ある調査の実施結果をIRTで分析すると、その調査に出題した「個々の問題の難しさを表す数値 (=困難度パラメータ。以下、困難度)」が算出されます。
- 調査X(出題した問題は、x1~x30の30問)と調査Y(出題した問題は、y1~y30の30問)を実施して、それぞれの結果をIRTで分析し、問題x1~x30の困難度と、問題y1~y30の困難度を、それぞれ算出できたとします。
- 問題x1~x30の困難度と、問題y1~y30の困難度は直接比較できません。それは、温度を表す「摂氏(°C)」と「華氏(°F)」において、0°Cと0°Fが異なることと似ています。ちなみに、摂氏と華氏には、 $^{\circ}\text{F} = 1.8 \times ^{\circ}\text{C} + 32$ という関係が成立します。
- 問題x1~x30の困難度と、問題y1~y30の困難度を互いに比較できるようにする手続きを「等化」といいます。等化のイメージを次のスライドにまとめました。



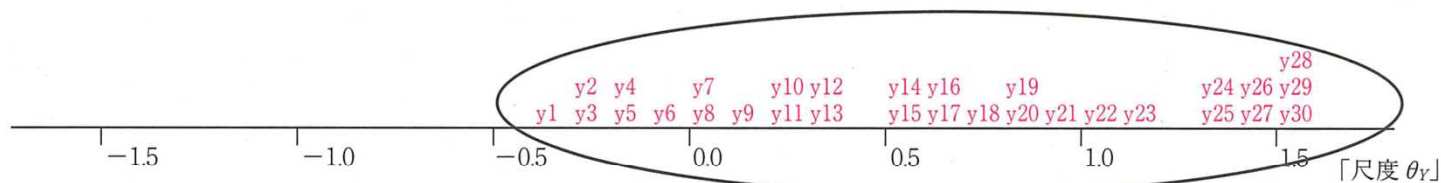
(参考) 等化のイメージ



調査 X に出題した問題の困難度



調査 Y に出題した問題の困難度



調査 Y に出題した問題の困難度を **変換** する

