測りすぎ　話題提供

2021 年 7 月 18 日　仲嶺真

1.　本書の概要（簡単な要旨 p.18-22 から）

近年，どこでも見られるようになった文化的パターン ＝ 測定基準への執着（測定執着）

：測定が実践されたときに意図せぬ好ましくない結果が生じるにもかかわらず，以下の信念が持続している状態

- 個人的経験と才能に基づいて行われる判断を，標準化されたデータ（測定基準）に基づく相対的実績という数値指標に置き換えるのが可能であり，望ましいという信念
- そのような測定基準を公開する（透明化する）ことで，組織が実際にその目的を達成していると保証できる（説明責任を果たしている）のだという信念
- それらの組織に属する人々への最善の動機づけは，測定実績に報酬や懲罰を紐づけることであり，報酬は金銭（能力給）または評判（ランキング）であるという信念

測定執着の基本的前提（なぜ起こる？）

：測定と改善の関係に関わる

- 測れないものは，改善できない　→　測定できるものはすべて改善できる
- 測定によってのみ組織が真に責任を果たせるようになると考える
  →　実績 ＝ 標準化された測定

⇒「透明性」が求められると可能な限り多くの情報を可視化・公開することが求められていると思う

：すべてが測定できるわけではなく，測定できることの大部分は重要ではない

- 複数の目的があるにも関わらず，特定の目的を過度に強調
  →　別の目的がとらえきれていない　→　もっと測定しよう　→　無用なデータと労力だけ増える

測定執着と伴って起こること，起こりやすいこと

：仕事の性質の有害な変化

- プロの職業倫理や判断と矛盾する可能性のある目標のせいで士気が低下
  →　指標の操作　→　結果的に組織にとっての機能障害へ

：実績に対する（ディス）インセンティブ

- 基本的には効果的でない。理念的なミッションを帯びた機関（教育，医療など）ではなおさら。
  →　数値目標への適応　→　イノベーション・独創性の抑制　→　短期目標の重視
  ※ 上の「指標の操作→」プロセスももちろん起こりうる

：問題に対する解決策が存在しない状況では，測定（収集と公表）自体が「解決策」として捉えられがち

- データを集めて公表するために費やした努力が精神的な熱意を満足させる
  ※ これは測定執着が起こる一因でもある（再帰性）

測定が効率的という信念はカルト的（みせかけの科学）

本書の目的
：実績基準が有益な場合の特定
：測定執着に特徴的な機能不全なしに，測定基準を活用する方法

2.　本書の結論（測定執着による悪影響は？）

「意図せぬ，だが予測可能な悪影響」p.172-176

- ☐　測定されるものに労力を割くことで，目標がずれる
- ☐　短期主義の促進
- ☐　従業員の時間にかかるコスト
- ☐　効用の逓減
- ☐　規則の滝
- ☐　運に報酬を与える
- ☐　リスクを取る勇気の阻害
- ☐　イノベーションの阻害
- ☐　協力と共通の目標の阻害
- ☐　仕事の劣化
- ☐　生産性のコスト

3.　いつどうやって測定基準を用いるべきか p.178-186

測定は判断の代わりにはならない。測定自体が判断を要する。測定を使うべか，どう使うかの確認として。

(1)　どういう種類の情報を測定しようと思っているのか？
(2)　情報はどのくらい有益なのか？
(3)　測定を増やすことはどれほど有益か？
(4)　標準化された測定に依存しないことで生じるコストはどんなものか？実績についてほかの情報源があるか？
(5)　測定はどのような目的のために使われるのか，言い換えるなら，その情報は誰に公開されるのか？
(6)　測定実績を得る際にかかるコストは？
(7)　組織のトップがなぜ測定実績を求めているのかきいてみる。
(8)　実績の測定方法は誰が，どのようにして開発したのか？
(9)　もっともすぐれた測定でさえ，汚職や目標のずれを生む恐れがあることを覚えておく。
(10)　ときには，何が可能かの限界を認識することが，叡智の始まりとなる場合もある。

妥当性の問題（ここでの妥当性は心理学における妥当性よりも広く取っている）
妥当性以前の問題（その測定に意義はありますか？という問題）
(3) 仲嶺・上條（2019）の「独自性」と関わる問題
(6) 研究倫理と関わる問題（調査＝測定が他者の負担になっていないか）
(10) 測定は万能じゃないよ，ちゃんと考えよう，という話
⇒ 測定の問題は，大きくは妥当性（その測定実践は適切ですか？）の問題

4.　妥当性とは

「妥当性とは，テストの使用目的に伴うテスト得点の解釈が証拠や理論によって支持される程度のこと」

（AERA, APA, & NCME, 2014）

Cf.「妥当性とは，テストの特性である。妥当なテストは測定しようとする属性の変動の結果を伝達する。」

（Borsboom，2005）

→ これはおそらく誤り。本書 p95-96「共通テストが生徒の知識や能力を測れなくなる」という事例。

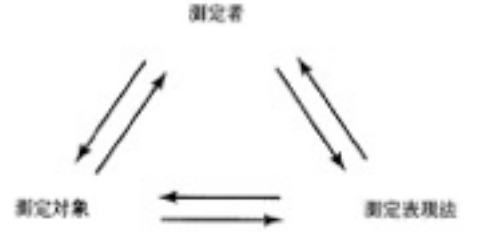

図 2.1 「測定者」，「測定対象」，「測定表現法」の相互規定の関係

吉野（2007）「公理的測定論の基礎」『数理心理学』p.66 より引用

現代の心理・教育的測定における妥当性は吉野（2007）が指摘する測定観に則っている？！

5.　妥当性のエビデンスの 5 つの源泉（AERA et al., 2014；平井，2016,）

・内容面のエビデンス：内容妥当性と呼ばれていたもの，実施手続き，採点手続きなど

・回答プロセス面でのエビデンス：回答中の思考プロセスの検討や，目の動き，反応時間など

・内的構造面でのエビデンス：因子分析の結果や信頼性係数など

・他の変数との関係性によるエビデンス：基準関連妥当性や収束的・弁別的妥当性と呼ばれていたもの，法則論ネットワークなど

・テスト実施の結果によるエビデンス：構成概念に関係のない部分で特定の集団に得点差が生じること，得点解釈の延長上に行われる主張（例，アチーブメントテストの結果から成績の振るわない学校を見つけるというシステムについて“テスト結果は，学習向上につながる”と主張すること）や，テスト実施により意図せざる影響が出ること（例，小論文テストをコンピューターで自動採点した結果，測定の一貫性は高まったが教師の採点スキルが低下したこと）など

→ 要するに，妥当性はプロセス（村山，2012）

6.　妥当な測定はどこにあるのか？

妥当な測定が可能となる根本的条件とは？→重要な示唆として（8）

- 「何か」を測定したいときに，第三者的に「何か」を把握することは可能なのか？
- 「現場」＝二人称的視点でしか「何か」を測定できないのではないか？

そもそもとして，心理学における測定とは，「測定表現の基盤/背景となる何か」を測っているのか？

- 吉野（2007）的測定観に則ると，測定対象とは，測定者と測定表現法とで「構成されるもの」では
- そうではないとしたら，なぜそう言えるのか

7. 全く別の問題として：透明性と統制性（コントロール性）

「測りすぎ」の背景にある大きな要因は以下の 2 つ？

- 透明性／客観性の問題　→　こっちは論じられた（透明性が実績の敵になるとき：p.162-167）
- 統制性の問題　→　こっちの扱いは不十分（(8)を論じるときに扱われていた問題であるけれども）

何かを「統制・コントロールしたい」という欲望から「測りすぎ」が生じている

- 研究も同じ欲望をもつ（自然をコントロールしたい，人間を制御したい，など）
- 「人間」理解の営みのはずが，「人間」統制の営みに（→測りすぎへ）
  ※ そもそも，理解と統制を切り分けられるのか？
- 営み自体も目的化（方法の目的化←『測りすぎ』でも指摘されていた）
- 測定（＝統計）と統制（=管理）とは切り離せないことの帰結（ハッキング『偶然を飼いならす』）

「統制の欲望」をどうやって手放すことができるのか？

- 統制の欲望を手放すよりも，統制の欲望とともに「偶然性への身投げ」も受け入れるという問題？
- だとしたら，相反する欲望をどのように共存・醸成する？

引用文献

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Borsboom, D. (2005). *Measuring the mind: Conceptual Issues in Contemporary Psychometrics*

平井 洋子（2016）．妥当性理論の歴史的変遷と心理学研究への適用に関する一考察――*Standards* を中心に――　首都大学東京人文学報，*512*，15-26.

村山 航（2012）．妥当性――概念の歴史的変遷と心理測定学的観点からの考察――

仲嶺 真・上條 菜美子（2019）．「心理学研究」の新心理尺度作成論文に記載された尺度作成の必要性　心理学研究，*90*(2)，147-155.

吉野 諒三・千野 直仁・山岸 候彦（2007）．数理心理学――心理表現の論理と実際――