

作題FD

-試験問題作題の基礎基本-

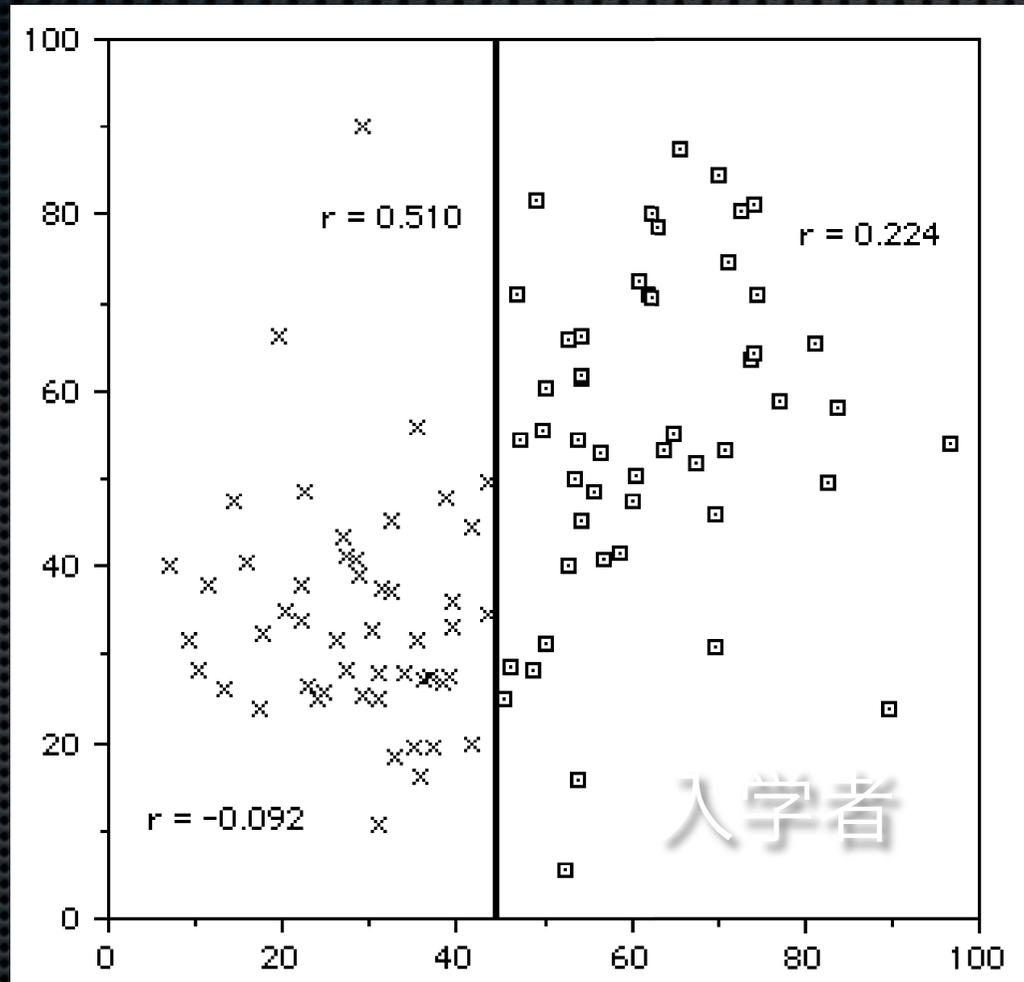
アドミッションセンター 吉村宰

作題手順の要点

- テストについての共通理解：「テスト＝測定用具」
- 仕様を策定する
 - 何のために、誰の、何を測定するか
 - どのように測定するか
 - 管理，結果の利用・分析，情報公開，ミスの防止など
- 品質のチェック（→次回の改善へ）
 - 妥当性：測るべきものが測られているか→追跡調査等
 - 信頼性：少ない誤差で測定されているか→テストデータの事後分析

選抜効果（群の切断効果）

大学
学業
成績



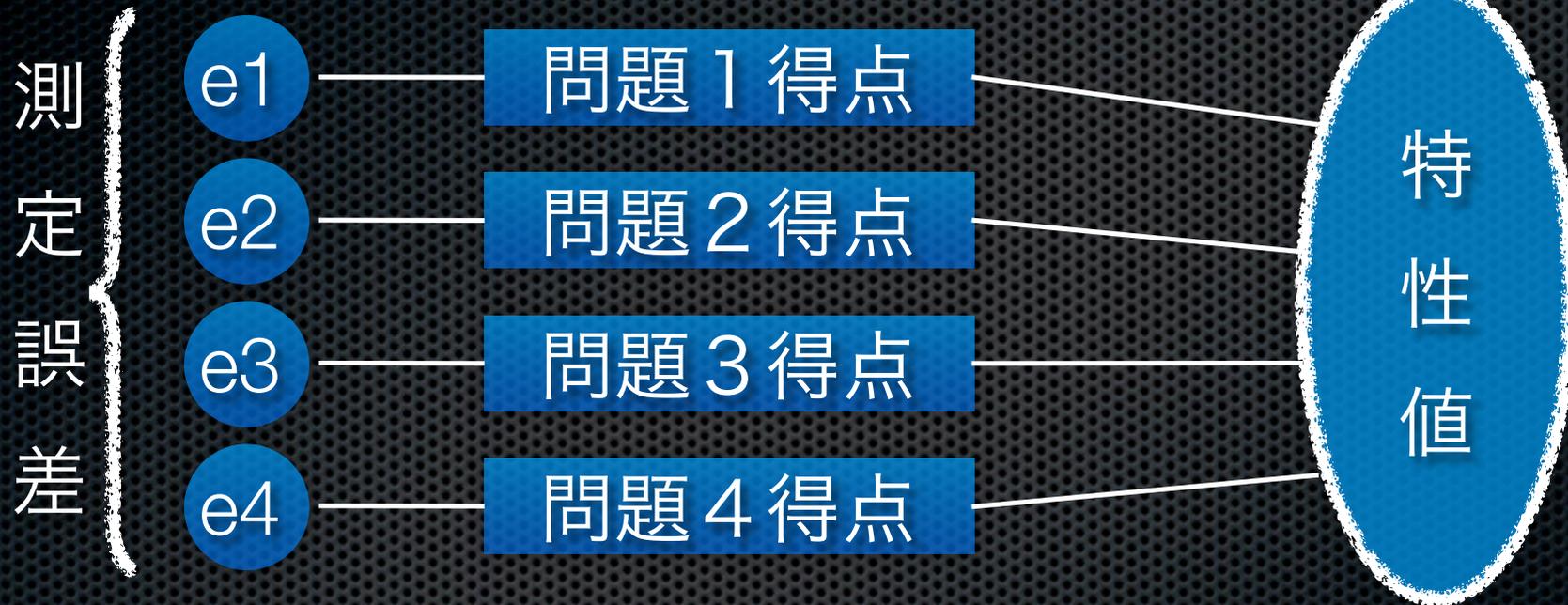
入試成績

テストとは

1. テストは測定の実具である

- 個人あるいは集団の、学力、能力、資質、態度等の特性を測定するための実具
- テスト得点は、何らかの特性の量の大小、多少、有無を表すものである
 - 好みの表明ではない
 - 生徒や学生への褒美ではない
(ボーナス点、出席点、努力点. . .)
- $\text{テスト得点} = \text{特性の本当の値} + \text{測定誤差}$

テスト（試験）：特性値の測定



仮定（モデル）

高いテスト得点となる

特性値が高ければ

高いテスト得点である

高い特性値をもつ

なぜ誤差？

- 出題しうる問題の母集団はほぼ無限
- 全領域から出題することすらできない
- 出題される問題のセットは問題全体のほんの一部
- 同一の受験生でも、問題セットが異なればテスト得点
が変化する ⇒ 確率誤差として扱う
- 誤差を考慮して、テスト得点はある程度おおまかに
扱った方が良い

2. テストには品質がある

- 妥当性：測定したいものを測定しているか
 - 専門家による問題内容の点検：教員による話し合い
 - 外部基準との相関の高さ：入学した学生のその後を観察
- 信頼性：測定の誤差は小さいか
 - 信頼性係数（ α 係数）：テストデータの事後的解析
 - 項目間，評定者間相関：テストデータの事後的解析
- テストの評価を行い次のテストの改善につなげる

大学入試

- 入学者の選抜に利用するテストである
 - センター試験や教科の学力検査だけでなく、総合問題、小論文、面接、志望理由書もテストである
 - 我々は単なるテストの利用者でなく、設計、作問、構成、実施、採点、管理などテストに関するすべての過程に携わっている
 - 入試は受験生の社会的処遇を決めるものであり、そのすべての過程についての説明責任を負う
 - 品質の高い入試を行う責務

3. テストの基本設計：目的と測定内容

- 目的

- アドミッションポリシーに沿った入学者を選抜する

- 測定内容

- 何を測るか：どのような学生を入学させたいかに依存
- 求める特性：具体的な行動レベルで記述する
 - 「～～力」という表現で終わってしまっはいけない
- どのようなことができたいかをリストアップ
- 在学生を調査できればなおよい

3. テストの基本設計：測定形式

- 測定したい特性をどうやって測定するか
 - 問題の内容
 - 問題の形式
 - 文字，図，音声，画像による問題の提示方法
 - 口述式，記述式，選択式などの回答形式，選択肢数
 - 問題数，配点，回答時間
 - パフォーマンステスト
 - 問題の形式も妥当性，信頼性に影響する

パフォーマンスに基づく評価の例（体育）

- クラスでの試合を実況する
- 架空の試合の実況台本を作る
- 試合の審判をする
- 同級生のダンス（実際あるいは映像）を批評する
- クラスターナメントのレポートを学校新聞に書く
- チームのコーチをする

3. テストの基本設計：「テスト全体」

- 「テスト全体」という観点
 - 質問項目の内容，測定形式，回答時間，回答の収集方法，採点方法，採点結果の利用，質問や回答の分析，管理，各種情報の公開までを見通した全体的な設計（仕様書としてまとめておく
とよい）
- ミスの防止
 - テストの全過程それぞれにおいて，実施者，回答者のミスを最小限にする努力
- 結果の利用方法
 - あらかじめ明記しておく（募集要項に記載済み）

小論文試験の作題と採点

1. 小論文試験の特徴

▪ 論述式

- 「知識や単純な技能」以外の特性を測定することができる
- テーマについての知識・関心・考え，書くことによって人に考えを伝えることができるかどうかを知ることができる

▪ 客観的な正解がない→主観的評定

- 明確な採点基準，採点手続きを準備する必要がある

2. 主観的評定の留意点

- 「複数の採点者＝様々な観点」⇒NG！
 - 採点者の好みによる選抜（AO入試の現状）
 - 何を測定したのか不明＝入学者の特性が不明
- 問題の作成と採点の工夫が必要
 - 一意に解釈できる課題やテーマ，指示の設定
 - 採点のためのスケールを作成（＝ルーズブリック）
 - どの段階に当てはまるかの判断にズレがあってもなにを判断するかに違いがあってはならない。
 - 十分な事前，事後の打ち合わせ

ルーブリック

- 採点基準表（資料参照）
 - 包括的ルーブリック，分析的ルーブリック
 - 入試には包括的ルーブリックの方が適している（と思う）
 - 各観点に評定者を配置して分析的ルーブリックを用いるのもよい（かも）
 - 採点者は少数でよい
 - 事前の訓練，事後のチェックが望ましい
 - 入学者に何を求めるかが明確になっていなければ作成できない。

3. 作題の留意点

- 受験生に何をどうすればよいかを明確に伝わるか？
 - 「～～あるいは××しなさい」：NG
- その課題をさせることで重要な特性を知ることができるか？
 - 何を知りたいかを事前に十分に話し合う必要がある
- 標準的な生活水準で十分に身に付く特性か？
 - 特殊な技能や経験を求めるのは不適切

面接試験の作題と採点

1. 面接試験の特徴

- 日本では面接試験に関する研究がほぼ皆無
 - 企業の採用面接研究：米では抱負（80年くらいの歴史）
 - 大学入試での面接試験に関する研究は米でもほぼ皆無
- 一般的に信じられている面接のメリット
 - 志願者の受け入れの決定を組織のメンバーによって行える
 - 優れた資質の志願者を選抜できる
 - 志願者の様々な知識，技能，能力（KSA）を効果的，実践的に測定できる（パフォーマンスに基づく評価）

2. 面接試験の欠点

- 信頼性が低い
→再現性がない
- 妥当性が低い
→あてにならない
- 有用性が低い
→使えない

面接者が犯しやすい過ち①

- 優れた点(悪い点)を1つ発見すると、他の全ての面も優れている(悪い)ように見てしまう
 - 学生時代に体育会系のキャプテンをしてきた応募者をいったん行動力がありそうだと評価すると、性格も良く判断力や計画性などなども兼ね備えた素晴らしい人物であると評価してしまう
- 面接者の好み、独自の価値観で人物評価をしてしまう
 - 自分の好みの服装や容姿の人物、また同じ趣味を持った人物を実際以上に高く評価する
 - 自分と異なる意見を持つ人物を低く評価する

面接者が犯しやすい過ち②

- 他の応募者との対比で、ある特徴を必要以上に強調して見てしまう（面接者自身の特性との対比でも起こる）
 - 明るく活発な応募者の次に会う人物を、実際以上に元気がないと判断する
 - 何事にも非常に意欲的に取り組む面接者が、自分との対比から応募者を実際以上に意欲がない人物と判断する

面接者が犯しやすい過ち③

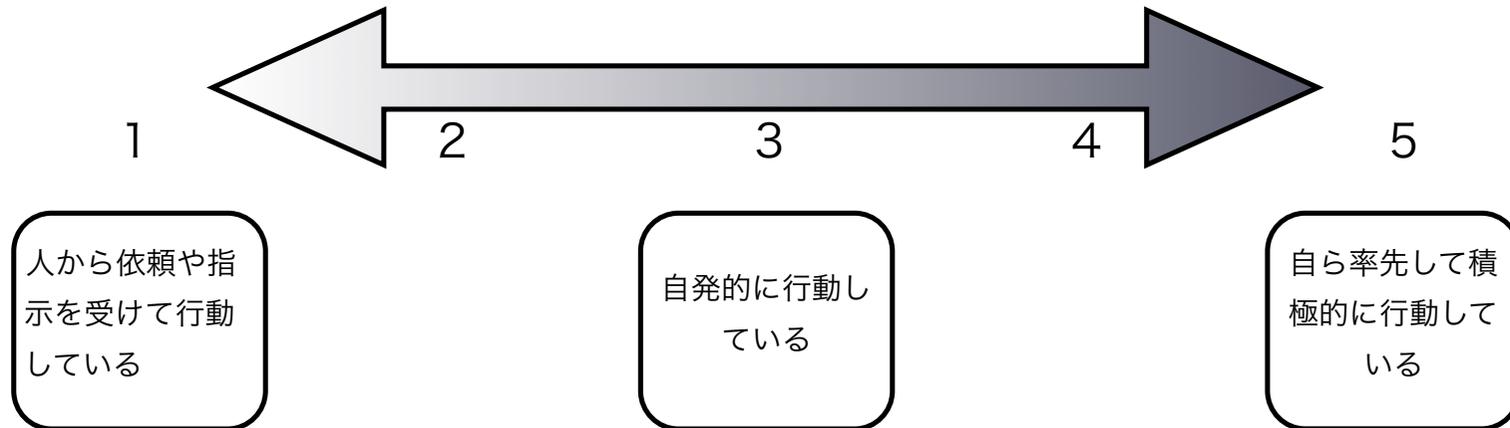
- 面接者がしゃべりすぎる
- 志願者それぞれに違う質問をする
- APとは無関係（ほとんど無関係）の質問をする
- 志願者をリラックスさせない（緊張させる）
- 評価することに関して過度の自信を持っている
- ステレオタイプや個人的なバイアスを持っている
- 志願者の非言語行動に影響される
- 誰に対してもほとんど同じ評価しかつけない
- 最初の数分間で評価をする

3. 構造化面接

- 適切に設計され，正しく行われた構造化面接（半構造化面接）であれば実用に耐えうる
 - （例）1人の面接者による構造化された面接の方が，構造化されていない複数の面接者による面接より妥当性が高い
- あらかじめ面接で評価する要件と評価基準を明らかにし，評価するための質問も事前にある程度設定したうえで進める形式の面接
 - 構造化面接は，測定手法としての面接の信頼性，妥当性を高めることを目的としている

要件：自主性

<評価基準>



< Main Question >

高校時代の活動で、自分から進んで周囲に働きかけて何かを始めたことがありましたか？そのときのことについて詳しく聞かせてください

< Sub Question >

どういった経緯でそれを始めたのですか

なぜ自分がやろうと思ったのですか

どのような気持ちで取り組んでいましたか

あなたが実際にとった行動をなるべく具体的に話してください

構造化面接の一般的な特徴（入試の場合）

- 志願者に共通の質問を行う
- APに基づいて面接の質問と評価方法を決める
- 具体的行動を表示した評価スケールを使う
- 特性ごとの評価を行った後、総合評価を行う
- ひとりの志願者に対して複数で面接を行う
- 面接者は事前に訓練を受ける

4. 構造化面接の設計のポイント

- 見るべき資質等を学生生活における具体的行動レベルでイメージし，面接要件として定義する
 - 評価する項目は，面接で評価することにメリットがあるものを中心にする（対人関係能力，態度的なものなど）
 - 評価要件はなるべく具体的に記述する
 - 「積極的な人」
→ 「対処すべき問題が起こったときに自ら率先して行動に移す人」

質問の準備

- 面接要件を評価するために妥当な面接質問を設定する
(妥当かどうかは議論で検討する)
 - 面接30分あたりの評価要件は2~3個まで
 - 1つの評価要件について複数の質問項目を用意する
 - 事前情報は原則として利用しないひとつの側面を評価するために、可能であれば異なる場面での経験や様子を尋ねる。ただしそのために話題が次々と変わるようでは逆効果となる。
 - 意見や考えを尋ねるのか、過去の経験を尋ねるのかの違い
 - 過去の行動を尋ねる場合は、その時の状況が想像できるぐらい事実情報を得ることが理想的

質問の種類

- 状況想定型質問

- 今日の授業で発表することになっているが風邪で高熱が出てしまった。大学に入ったばかりで知り合いもいない。あなたはこの状況でどうしますか？

- 行動記述型質問

- 大学ではグループで共同作業をすることがよくあります。最も最近、あなたがグループの一員として働いた経験について教えてくださいませんか？

スコアリングの準備

- 体系だったスコアリングを行う
 - 個々の質問, 評価項目, 総合評価の3つのレベル
 - 評定用の尺度を準備する
 - 評定のアンカーはなるべく行動で定義
 - 評価を支援する面接評定票を作り込む

面接者の準備

- 面接者の打合せにおいて、求める人物像と評定の指標との関連を確認する
- 面接者の基本的な面接スキルの向上をはかる
 - 面接者が犯しがちな過ちの自覚
 - 聞いてはいけないこと、差別的な発言

面接者の心構え

- 志願者とのやりとりをスムーズにする
- オープンなやりとりができる雰囲気をつくる
- 一貫した質問を行う
- 人の話を聞く
- 適切なメモをとる
- 誘導的，高圧的な態度は避ける
- 非言語的なてがかりを制御する（惑わされない）
- 自身の面接観にこだわらない

参考文献・参考資料

- Handbook of Test Development, Dowing, S.M., Halyadyna, T.M. (Eds.), LEA, 2006.
- Human Resource Selection, Gatewood, R.D., Feild, H.S, South-Western, 2001.
- 今城志保, 入試における面接の理論と方法—採用面接の現場に学ぶ—, 第1回長崎大学アドミッションセンターFD資料, 2007.12.7
- 中畝菜穂子, 選抜テストとしての面接—人事測定場面での面接改善の試み—, 2007年度長崎大学アドミッションセンター研究会資料, 2008.2.19