

新SATテストについて

— 改訂の背景ならびに改訂点を中心として—

石岡恒憲・鈴木 規夫(大学入試センター)

SAT Reasoning Testとは何か

- アメリカの教育団体であるCollege Boardが主催・運営
- 4年制大学への進学希望者を対象に行われる全国共通試験
- 大学での学習についていけるだけの論理的思考力
- SAT I (Reasoning Test: 問題解決型テスト)
- 1994年 Achievement Test → SAT II (Subject Tests: 科目別テスト)
- 2004年 数字のIとIIが脱落; SAT Reasoning Test, SAT Subject Tests
- 2005年3月 新SAT導入;大学の授業内容により即した試験内容に変更

目次

- 新SATの概要
- 今回の改訂の理由
- 改訂の基礎となる事前研究
- エッセイ試験
- 今後のSAT

新SATの概要

- Reasoning Test
- Subject Tests

Reasoning Test

- 3時間45分
 - 「問題解決」,「論理的思考力」
- (1) Critical Reading (従来のVerbal; 75分→70分)
 - (2) Writing (60分)
 - (3) Math (75分→70分)

Critical Reading (70分)

- Sentence Completions (文章完成)と Passage-based Reading (読解)
- Analogies (類推)の削除

Writing (60分)

- 新セクション
- Multiple-Choice (多肢選択;35分)とShort Essay (エッセー;25分)
- 文章構造や文法、語法などの間違い探し
- 単文・エッセーの主旨を読み取って答える
- Short Essay :与えられたトピックについて自分の見解をまとめる

Math (70分)

- Statistics (統計学)、Algebra (代数)、Geometry (幾何)など
- Multiple-choice (多肢選択問題) + grid-in (数字 /.) 10問
 - 計算機を持ち込み可
- Quantitative Comparisons (数量比較) の削除
- Algebra II, 指数成長 (exponential growth)、絶対値、関数表記、負や分数の指数、の追加

Analogies (類推) の問題

- ・ 大文字で示されている単語あるいはフレーズのペアと最も似た関係のあるペアを選べ:

1. CRITERION : JUDGEMENT ::

- A. song : voice
- B. depth : bottom
- C. standard : selection
- D. edge : sharpness
- E. height : statue

Quantitative Comparison (数量比較)の問題

- ・ 数量を比較しカラムAが大きければAを、カラムBが大きければBを、等しければCを、不定ならばDを選べ。

1. 50の n パーセントが40より大きいとき::

カラムA : n

カラムB : 90

Subject Tests

- 英語(つまり国語)、歴史・社会学、数学、自然科学、語学の5分野
- 英米文学、アメリカ史、世界史、数学(レベル1、レベル2)、生物学、化学、物理
- リーディング問題のみの語学(フランス語、ドイツ語、近代ヘブライ語、イタリア語、ラテン語、スペイン語)
- リスニング問題を含む語学(中国語、フランス語、ドイツ語、日本語、韓国語、スペイン語)

Subject Tests (Cont.)

- 受験科目は志望する大学・学部による
- 予め指定する大学/出願者が選べる大学/
スコアを求めない大学
- テスト時間:1時間/科目、形式:多肢選択
- Writing Subject Test 削除(Writing
Sectionに移行)

今回の改訂の理由

- College Boardの言い分
- ブルーリボン委員会

College Boardの言い分

- ・ 受験生が高校時代に学習してきたことをより反映させるため
- Writing: 大学や将来において成功するための重要なスキル
- Math 代数II の追加
 - 高校生の70%が卒業前年の終りまでに代数IIを履修
 - 進学希望者の97%が3 years of math (9-12年生の3年次の数学)を終了
- Analogies削除: 高校のカリキュラムと直接的に結び付いていない

カリフォルニア大学による働きかけ

- Analogies削除、Writingを加えるべきという助言
- 旧SAT + Writing Subject Test → 全体的な予測妥当性 .06-.08上昇
- Analogies : 出題形式が語彙の単純暗記を促進

ブルーリボン委員会

- CBもカリフォルニア大学の助言に注目してきたことを認めている
- 改訂は全てのメンバーからのフィードバックを考慮
- 他の多くの要因も変更に寄与
- 今回の改訂の種: 1990年 ブルーリボン委員会による報告書 “Beyond Prediction (予測を越えて)”

ブルーリボン委員会 (Cont.)

- 1994年 改訂はこれに基づく
 - 反意語antonymの問題が削除
 - より長文の読解が追加
 - 数学:Grid-In追加、電卓の使用
- Writing Sectionを加えるべきとの提言
 - 検討されたものの採択には至らず;今回はじめて導入
 - 技術的な処理能力の問題;百万もの受験生のエッセイをプロの評価者へ転送することができなかった
 - 現在 Peason Educational Management社のインターネットに基づいた採点システム ePen

改訂の基礎となる事前研究

- 新SAT：2002年6月 アナウンス
- 3つの段階
 - フィールド試行テスト以前
 - フィールド試行テスト
 - フィールド試行テスト以後

試行テスト以前の研究

- (1) 信頼性確保の問題
- (2) 疲労の影響
- (3) エッセイのタイプ

(1) 信頼性確保の問題

- Critical Readingの変更と Analogies問題の削除
- 変更/削除→測定精度の信頼性を保証
- 実際のSATのデータを用いた検証
- 信頼性維持は可能
- アイテム困難度の分布(両裾)を修正する;
- 非常に易しい問題と、非常に難しい問題を増やす必要

(2) 疲労の影響

- 新SATでは試験時間が45分増える
- テストにおける疲労を 100人の学生を用いて調査
- 読解および数学の試験においては 5時間から6時間で疲労に耐えられなくなる

(3) エッセイのタイプ

- ・ エッセイ試験でどのようなタイプの質問文にするか
- ・ 人種、母語、性差の違いで成績の違いが生じるか
- ・ 試作問題は特定のグループに不利であるとは認められない

試行テスト

(1) 目的

(2) 実験計画

(3) 結果

(1) 目的

- 2003年3月フィールド試行テスト
- 新SATの周知
- 内容、統計的性質、時間配分の評価
- 新SATスコアが当時のSATスコアと比較可能か

(2) 実験計画

- 679の高校から45,000人以上の学生
- 公立/私立、都会/地方、全米の全ての地理的地域
- 人種、民族の違いに基づく研究結果を保証
- アフリカ系アメリカ人/スペイン系(ヒスパニック)/ラテン系の学生を高比率で募集

(3) 結果

- 新しいCritical ReadingとMathのセクションはテストの困難度に影響を与えない
- 新SATの信頼度は現在(当時)のそれとほとんど同じ
- 新SATスコアと当時のそれとの相関は、3つの全てのセッションにおいて極めて高い(.95-.97)
- 性や民族の違いによるスコアの差を悪くすることはない

試行テスト後の研究

- (1) 制限時間
- (2) 問題配置
- (3) 予測妥当性

(1) 制限時間

- 制限時間を様々に変えて、学生が答えることのできる問題数との関係
- 全設問の75%が解答され、被験者の80%が最後の問題に到達するという時間制限が最も適当
- テスト信頼性を上げるために、Writingのセクションを10分多くすることが決定

(2) 問題配置

- 大半の学生はテストの最初にエッセイを選択する傾向
- 試験の最初にエッセイ試験をおいた方が、少しスコアがよくなる
- 新SATではエッセイは常に最初に置かれる

(3) 予測的妥当性

- Writingセクションを追加することの最終的な目的
- 大学での成功を予測するための妥当性向上
- American Institutes for Research社: 13大学
1,200名
- Writingセクション追加 → 成績の予測妥当性(相関係数).01-.02増加

→Writingセクションの追加の価値

エッセイについて

(1) 評価の視点

(2) 解答時間

(1) 評価の視点

- テーマに対して効果的かつ洞察力をもった意見の展開
- 批判や例示等による補助的な情報提示
- 組織化、焦点が明確、アイディアの挿入がスムーズ
- 言語の使用が効果的、語彙が多い
- 文脈の構成が意味を持ちながら多様
- 文法上の誤りがない

(2) 解答時間

- わずか25分(ワード数の制限なし)
- 十分に推敲されていることを期待していない
- スペリング・エラーも理解に支障がない限り影響を与えない
- 文字の巧拙も評価しない。読み手の判読は必要
- コンピュータを使用すれば解決する作文上の問題には便宜

今後のSAT

- 新テストの調査；技術的な特徴がうまく機能しているか
- どのような学生層にどのような影響があるか
- 新たな妥当性研究→ 約10年後のSATが決定

ご清聴ありがとうございました

FAQ for SAT

スコア

- 200-800のスケール;200が最低で800が最高;等化後のスコア
- 受験生のスコアは素点で計算
 - 正解すれば1点
 - 多肢選択において不正解には-1/4点
 - Grid-Inにおける誤答/無解答は減点されない
- 最終的な素点は小数点以下を四捨五入;その後、等化
- スコア幅は上30点、下30点;ごく短い間に再試験したら確率50%でこの幅に入る
- Critical Reasoning/Math:パーセンタイルで20%、スコアで60点以上の違いは差がある

スコア(Cont.)

- Reasoning Test: 各セクションで200～800
- Writing セクション:多肢選択問題を20～80、エッセーを2～12
- エッセーのサブスコアは各大学で扱いが異なる
- Subject Tests : 各科目ごとに200～800
- 語学:サブスコアがレポートされるものもある

スコア通知

- ・ オンライン(要登録)/電話/郵送の3方法
- ・ オンライン/電話は受験日から約2週間後
- ・ 郵送は約3週間後
- 4校までは無料(受験料に含まれる)
- ・ 受験時に指定した希望大学へ College Board 本部から直接送付
- ・ オンラインの場合
 - － 無料で24時間確認
 - － 正答・誤答・無回答数、素点、スコア履歴などの確認

ETSとCBとの関係

- ETSがSAT/Advanced Placement Programを開発/
管理
- CBはこれらテストの引受人
 - テストの構成/運営の仕方/使用に対して決定権
- ETSからみればCBは大きな顧客
 - とともに多くの機関や委員会のために似た仕事

受験者

- SAT Critical Reading 延べ180万人(117万人)
- SAT Subject Tests 80万人
- Writing 21万人

試験日

試験日	テスト	US 登録日	海外登録日
10/8, 2005	SAT & Subject Tests	9/14, 2005	9/7, 2005
11/5, 2005	SAT & Subject Tests	10/12, 2005	9/30, 2005
12/3, 2005	SAT & Subject Tests	11/9, 2005	10/28, 2005
1/28, 2006	SAT & Subject Tests	1/4, 2006	12/22, 2005
4/1, 2006*	SAT Only	3/8, 2006	N/A
5/6, 2006	SAT & Subject Tests	4/12, 2006	4/3, 2006
6/3, 2006	SAT & Subject Tests	5/10, 2006	4/28, 2006

*SAT のみが US 国内及び準州、プエルトリコで実施

エッセイ問題例

- ・「人は富や名声よりも個人的な成功をなしとげることに
より動機付けされる。この問題についてあなたの視点を
掘り下げたエッセイを計画し、記述しなさい。理由付け
や読書、学習、経験、観察した例を示しながら自分の
見解を支援しなさい。」

エッセイ 採点の仕方

- 採点者: 大卒以上の学位;最低3年の教育経験;採点のための教育
- 時給\$17.00 ~ \$22.00 APやSAT Subject Testでの採点経験考慮
- 自宅のIBM PC (Windows); AM6:00-PM12:00
- 2人の評価者が互いに独立に評定
- 1~6点(6点が最高)/人
- 2人の評点が2点を超えて異なるときscoring leaderが調整
- 得点はWritingセクション全体の1/3