

「思考力・判断力・表現力」を評価する CBTを用いた情報科大学入試の試行

大阪学院大学教授 西田 知博

1. はじめに

2016～2018年度に大阪大学、東京大学、情報処理学会は文部科学省大学入学者選抜改革推進委託事業『情報学的アプローチによる「情報科」大学入学者選抜における評価手法の研究開発』において、「思考力 (Thinking) ・判断力 (Judgement) ・表現力 (Expression)」(以下、TJEと略す) を評価する情報科の入学者選抜に関する研究開発を行ってきた^[1]。この中で、情報科入試のためのCBT (Computer Based Testing) システムの開発と、それを用いた試行試験を行った。2017年度は情報処理学会情報入試委員会が実施してきた紙ベースの情報入試模試^[2]をCBTで実施できるように開発したシステム (V1) を用いて、大学生176名、高校生1406名を対象に試行試験を実施した^[3]。2018年度はV1をベースにCBTならではの機能を備えたシステムV2を開発し、それを用いた試行試験を実施した。ここでは、V2システムを用いた試行試験とその結果について紹介する。

2. CBTシステムV2

CBTシステムV2では、IRT (項目反応理論) を用いた出題が可能となるように、用意した問題プールからのランダムな出題を可能とした。

また、設問を一度に見せるのではなく指定された遷移にしたがって提示し、対話的な問いかけの中でトラブルシューティングやプログラムのテストなどを行わせ、そのプロセスを評価する出題の枠組みを実現するための機能 (ゲームブック形式) も追加した。プログラミングやアルゴリズムを問う問題は、V1では図1のような短冊を並べ

て解答する形式のみで出題したが、V2では図2のようなブロックプログラミングエディタを作成し、解答したプログラムや手順を実際に動作させ、その結果を確認できるような出題を可能とした。このような出題の枠組みでは、単なる最終解答だけではなく、解答の過程も見ることにより、TJEの評価ができる。

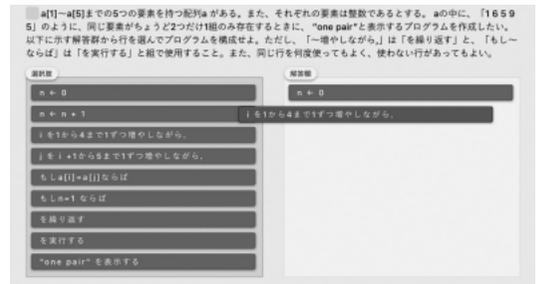


図1 短冊型プログラミング問題



図2 ブロック型プログラミング問題

3. 2018年度試行試験

2018年度試行試験 (以下V2試験) では、出題範囲を「データの分析・表現」と「プログラミン

「プログラミング／アルゴリズム」とし、IRTを想定した小間で構成したセット1と、センター試験などと同じ大間で構成したセット2の二つの問題群を作成した。

セット1は、複数の受験機会を提供できるよう、多くのCBTで実施しているような受験者によって問題がランダムに変わる試験の実施を想定したものである。また、そのような形式で出題する小間でTJEを問うことができるかを確認することも目的としたため、用語を問うような知識問題は出題しないこととした。

今回は約100問の問題のプールを用意し、それらを類似の問題で区分けし、それぞれからランダムに15問を出題するようにした。試験時間は30分である。「プログラミング／アルゴリズム」の分野の問題はV1と同様の短冊並び替えのプログラミング問題(図1)の他、プログラムのパラメータを選択させるもの(図3)、数直線上のロボットの動きを考えさせるもの(図4)などさまざまな形態で出題した。また、コインで動くコマの動きをトレースさせる問題(図5)や、データ表現を問うものとしてビット操作を選択式のものも含めた空所補充で答えさせる問題(図6)、データの分析を考えさせるものとして文章から判断できる事実を選ばせる統計の問題などを出題した。

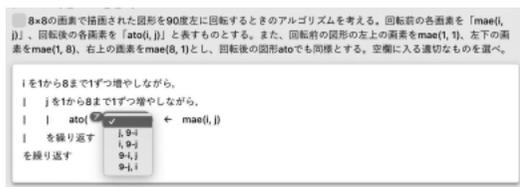


図3 パラメータを選択解答させる問題

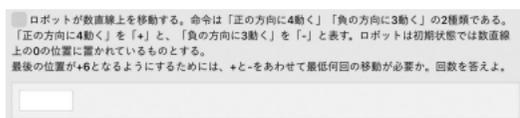


図4 ロボットの動きを考えさせる問題

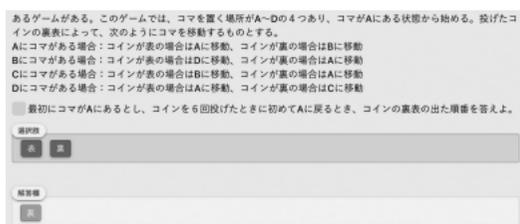


図5 コマの動きをトレースする問題

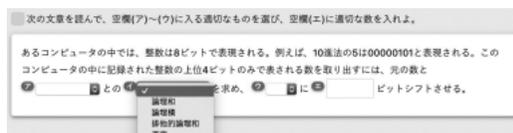


図6 ビット操作に関する問題

セット2は、「データの分析・表現」と「プログラミング／アルゴリズム」の分野の大問を1問ずつ、計2題出題するセット2種類(セット2a, セット2bと呼ぶ)を作成した。なお、どちらのセットを受験するかは、出題者側で割り当てを決めた。

セット2aの第1問(2a-1と表記、以下同様)は、ゲームブック型の問題で、新しく導入したレジの動作を、プログラムが見えない状態で検査(ブラックボックステスト)を行い確認するという問題である。購入する商品を選び、その消費税計算などが正しくされているかをゲームブック形式で検証していく(図7)。解答は、必要と考えた検証が終わった後、選択肢の中から該当する不具合を選ぶ形で行う。この問題は分野としてはプログラミングであるが、プログラムを作成するのではなくテストするという、これまでとは違った出題としている。

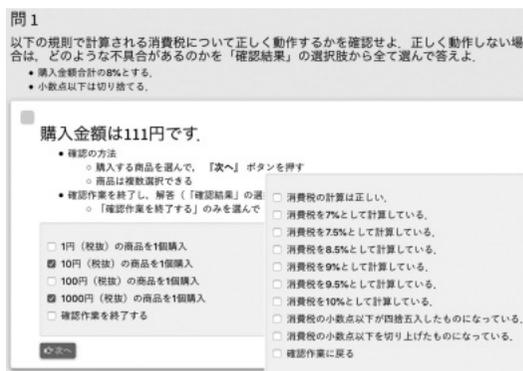


図7 ゲームブック型問題

2a-2はデータを扱うもので、自治会館の部屋の予約方法の改善を問う問題である。元の予約表の問題点を選ぶものや、予約受付プログラムを作るのであれば何をどのように表示したらよいかということ短冊並び替えのインタフェースを使って答える問題(図8)などを出題した。

2b-1はデータベースに関する問題である(図9)。生徒、部活、科目、成績という4つの表について、選択や射影や結合といった操作をすると

いう問題で、データベース操作などを日本語表記のブロックとし、それらを並び替えて解答する。データベース操作に関しては前提知識がないことを想定し、すべてを問題中で例示して説明した。

■(2)「予約のさまり」と「予約受付の入力ルール」にもとづいて、パソコンで受付処理を管理するためのプログラムを作成することにした。そこで、さらに下記の<条件>を設定した。受付担当者が、部屋の空を確認して、予約受付に必要な項目を入力し、確定する場合、表示する画面の流れはどのようにすればよいか。<手順>の項目を<画面の流れ>になるよう、ドラッグしなさい。ただし、すべての画面は使用しなくてよい。

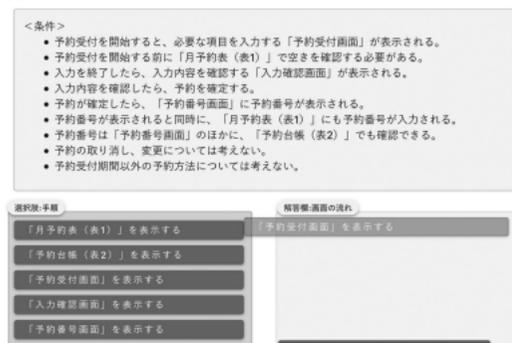


図8 部屋予約管理に関する問題



図9 データベース操作の問題

2b-2は、ロボットをマス目の中で動かすブロックプログラムの問題である。プログラムを読んでロボットが実際どのように動くかを答える問題と、指定された動作をするためのプログラムを図2に示したブロックプログラミングエディタで答える問題を用意した。プログラムはいつでも実行可能で、図2上のように実行結果が表示されるので、実際に動作を確かめながら解答することができる。

採点は、セット1は15問各1点で15点満点、セット2aと2bは第1問・第2問を各30点で、両セットとも合計60点満点で行った。

4. 試験結果

V1の試験と同様、大学1年生（大阪大学、東

京大学、計111名）と高校生を対象に試行試験を実施した。高校での試行試験は情報の授業を受けた1・2年生を対象とし、合計8校1531名が2019年1、2月に受験した。ただし、セット1と2を別の日程で受験してもよいとしたので、受験者数はセット1：1423名、セット2a：804名、セット2b：727名であった。各セットの得点分布は図10の通りで、平均点はセット1：4.69（標準偏差 2.27）、セット2a：22.87（標準偏差 13.55）、セット2b：11.05（標準偏差 9.86）となった。

図11はセット2の各大問の得点分布である。プログラミングに関連する2a-1と2b-2は比較的近い分布となった。「テスト」と「作成」という違う内容で出題したが、同等に能力を測れている可能性がうかがえる結果となった。また、ブロックプログラミングの問題2b-2は2割強が無得点であるが、情報入試模試など過去の短冊並び替えのプログラミング問題よりもその比率は少なく、実際に動作を確認できたことの効果がうかがえる。部屋予約管理の問題2a-2は平均と中央値が中間にあり、分散も広がり、正規分布に近い。現在の情報の授業では取り扱っていない内容であるにも関わらず、問題文を読み解く、間違いを探す、条件から必要な項目を選択するなど、思考力を働かせて解いたのではないかと推測する。情報の内容を広く捉えさせ、TJEを総合的に問う出題として、精査を進める意義のある問題となった。一方、データベースの操作を問う2b-1の結果は受験者727人中542人が無得点という結果になった。データベースの一般的な操作を日本語表記し、例を見てもらうことで初見でも解くことができると考えていたが、プログラミングと同等に操作結果を検証できるようにするなどの配慮が必要であったと考える。

セット1の得点との相関係数は2a-1：0.441、2a-2：0.452、2b-1：0.346、2b-2：0.507となり一定の相関は見られ、小問においても思考力などが測れる可能性があるという結果となった。しかし、明確な相関は確認できなかったため、更に多くの問題を試し、相関の有無を見極める必要がある。

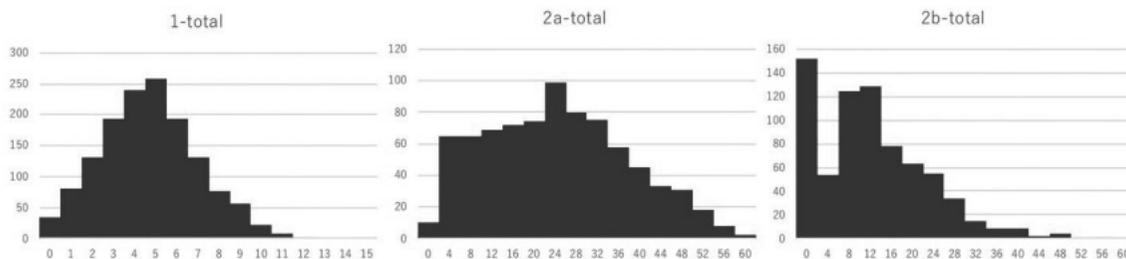


図10 各セットの得点分布

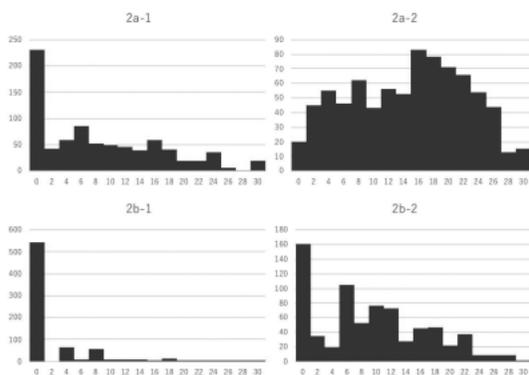


図11 セット2各問題の得点分布

5. アンケート

試行試験の受験者に対して試験後に操作性等についてのアンケートを行った。

「操作がスムーズにできたかどうか」という問いに対しては、セット2aでは「非常に操作しやすかった」「操作しやすかった」の合計が87.5%であったのに対して、セット2bは65.7%とやや低い結果となった。しかし、データベースの短冊並び替えやブロックプログラミングといった複雑な操作に対して比較的に多くの高校生が対応できているという結果になった。

CBTと従来の紙ベースの試験（PBT）ではどちらがよいかという質問については、V1がほぼ半々だったのに対し、今回はCBTならではの問題を入れたこともあり、6割以上がCBTの方がよいと答える結果となった。

紙ベースの試験がよいという理由としては、

- ・紙で書かれているほうが、いちいちスクロールしなくていい
- ・CBTでは画面の位置のみしか見ることはでき

ないが、PBTでは紙全体を一度で見ることができるから

・操作がバグったりわからないところがあるからのような回答があり、全体の見通しや操作への不安などが挙げられた。

一方、CBTがよいという意見としては、

- ・選択肢の並び替えが分かりやすい
 - ・自分が作ったプログラムを実行できたから
 - ・書き直したり組み立てなおしたりしやすい
- のような回答があった。短冊並び替え、ブロックプログラミングなど、CBTならではの機能が評価されたことがうかがえた。

6. おわりに

本稿では2018年度に実施した「情報科」入試の試行試験とその結果を紹介した。CBTは情報科だけに留まらず他の教科でも紙ベースでは評価できない能力の評価ができると考える。また、IRTを用いて複数の受験機会を設け、TJEを総合的に問える可能性はうかがえた。一方で、入学試験でCBTを使うためには公平な環境での試験実施が求められるので、今後も研究を重ね評価を行っていきたい。

参考文献

- [1] 萩原兼一：2025年度高校教科「情報」の入試における思考力・判断力・表現力の評価の考え方と方法，じっきょう情報教育資料No.46，pp.1-8，(Apr.2018)
- [2] 谷聖一他：第3回・第4回大学情報入試全国模擬試験，SSS2016 情報処理学会情報教育シンポジウム論文集，pp.7-14，(Aug.2016)
- [3] 西田知博他：「情報科」の大学入試実施のためのCBTシステムの開発とその検討，大学入試研究ジャーナル第29号，pp.117-123，(Mar.2019)