

大学初年次学生向けの ICT 知識テストの解答分析と問題再考 Analysis of Answers to an ICT Knowledge Survey for First-year University Students and Reconsideration of the Questions

小俣 昌樹[†] 金子 大輔[‡] 石田 雪也[§] 吉川 雅修[†] 古賀 崇朗[¶]
Masaki Omata Daisuke Kaneko Yukiya Ishida Masanobu Yoshikawa Takaaki Koga

1. はじめに

2003 年度から高校に教科「情報」が新設されたこと[1], これまで複数あった情報科目が 2022 年度からひとつの共通必修履修科目「情報 I」となったこと[2], そして, 2025 年の大学入学共通テストから「情報」が追加されること[3]など, 高大接続における情報教育に関する学習効果の評価が必要となっている. そこで著者らは, 高校普通教科情報の範囲の知識を問う全 40 問の五肢択一式の「ICT 知識テスト」を作成し[4, 5], 著者らが所属している大学の初年次入学生を対象にこのテストを実施してきた[6, 7].

本論文では, 著者らが所属する 4 大学の 2016 年度から 2021 年度までの初年次学生約 17,000 人を対象に実施した「ICT 知識テスト」の解答結果の傾向を分析し, 設問を再考する. 分析において, 各設問における 6 年間全体の正答率を導出し, 正答率の高い 10 問と低い 10 問から, その設問内容の傾向を考察する. そして, この考察に基づく設問の再考において, 正答率の低い 20 問の各学生の成績とその学生の全 40 問の成績との相関を分析したり, 正答率の低い 20 問の解答パターンから全 40 問の成績を推定するニューラルネットワークを構築したりする. この分析および再考に基づく本論文の貢献は以下である.

- 全体正答率が約 20% から 40% 程度の正答率の低い設問を明らかにすることで, 高校普通教科情報の教育の効果として, 知識があまり定着していないと考えられるトピックの候補を明確にした.
- 正答率の低い 20 問の解答パターンから全 40 問の成績を決定率 0.891 の精度で推定できることより, 従来の半数の設問数でも各学生の高校の情報知識に関する成績を把握できるという効率の改善を示した.

2. ICT 知識テスト

2.1 テスト問題

本研究で使用した ICT 知識テストは, 情報活用の実践力, 情報の科学的な理解, 情報社会に参画する態度という, 情報教育の目標における 3 つの観点をもとに, その知識部分に関する合計 40 問の五肢択一式問題で構成されている[5]. 各観点に関する内訳は, 実践 12 問, 理解 15 問, 態度 13 問であり, 1 問正解を 1 点として, 計 40 点満点とする. 5 つの選択肢の内訳は, 正答 1 つ, 選択肢用誤答 3 つ, および自信をもって解答できない場合に選択するための「わからない」という選択肢である. この「わからない」を設けた理由は, 当て推量で選ぶことを避けるためである.

[†] 山梨大学 University of Yamanashi

[‡] 北星学園大学 Hokusei Gakuen University

[§] 公立千歳科学技術大学 Chitose Institute of Science and Technology

[¶] 佐賀大学 Saga University

2.2 テストの実施

この ICT 知識テストを実施した大学は, 著者らの所属する山梨大学, 北星学園大学, 公立千歳科学技術大学, および佐賀大学である. 実施対象者は, 各大学の初年次入学生である. 実施時期は, 各大学の入学時から初年次前期中間期くらいまでである. 実施方法は各大学によって異なっており, 紙のマークシートを学生に配布して自宅で解答して大学へ提出する方法, または, 大学の e-ラーニングシステムを使って授業内や自宅で解答する方法である.

3. 解答結果および分析

3.1 解答者数

表 1 に, 各年度の受験者総数および分析対象者数を示す. 受験者総数に年度ごとの違いがあるのは, 大学によっては, 入学者全員を対象として実施できた年度とできなかった年度があるためである. また, 受験者総数よりも分析対象者数が少なくなっている理由は, マークシートの読み取り不備などで全 40 問分の解答を得られなかった受験者を除いたためである. 以降の分析および再考では, この分析対象者全 17,049 人の解答を対象とする.

3.2 正答率の高い 10 問

表 2 に, 全分析対象者の全年度における正答率の高かった順の上位 10 問の正答率および設問とその選択肢を示す. この表から, 正答率の高い設問には, 多くの受験者の日常生活におけるごく一般的な語句や状況が多く見受けられる. たとえば, 問 37 の SNS, 問 34 の映像, 問 24 の Web サイト, 問 11 のパスワード, および問 6 の検索エンジンなどである.

3.3 正答率の低い 10 問

表 3 に, 全分析対象者の全年度における正答率の低かった順の下位 10 問の正答率および設問と選択肢を示す. この表から, 正答率の低い設問には, アルファベット 3 文字程度で略されるような専門用語が選択肢に含まれる設問が多く見受けられる. たとえば, 問 39 の URL や POS, 問 16 の

表 1 ICT 知識テストの受験者数

年度	受験者総数(人)	分析対象者数(人)
2016	3,072	2,964
2017	2,762	2,417
2018	3,235	3,090
2019	3,161	3,018
2020	3,373	3,253
2021	2,399	2,307
合計(人)	18,002	17,049

表 2 全体正答率の高い設問

正答率 (%)	問番号	問題文と選択肢 (「わからない」の選択肢を除く)
88.4	37	SNS (Social Networking Service) の一般的な特徴として、当てはまらないものはどれか。 人と人とのつながりをオンラインでサポート・促進する／一度発信した情報であっても完全に削除できる／現在の自分の行動や考えを簡単に発信できる／常に正しい情報を入手できるとは限らない
87.6	34	スポーツなどの状況を伝えるとき、一般に、最もわかりやすく伝えることのできる表現メディアはどれか。 画像／文字／音声／映像
87.3	24	Web サイトから信憑 (しんぴょう) 性のより高い情報を得るための工夫として、最も適切なものとはどれか。 検索エンジンが最初に見つけた Web サイトから情報を得る／複数の Web サイトからの情報を比較・照合する／人気の高い Web 掲示板から情報を得る／訪問回数の多い Web サイトから情報を得る
87.1	32	効果的なプレゼンテーションを行う工夫として、最も適切なものとはどれか。 スライドの音や背景を派手にする／全て明るい色で表示させる／聞き手の年齢や立場を考慮する／身振り手振りを使わない
86.0	11	パスワードの設定や管理に関する記述のうち、最も適切なものとはどれか。 管理者から要求があっても教えない／文字数はなるべく少なくする／氏名や生年月日をパスワードにする／忘れないようにメモしてパソコンのモニタに貼っておく
85.3	21	明朝体やゴシック体など、文字の書体のことを何とというか。 ポイント／レイアウト／文字飾り／フォント
81.5	6	検索エンジンを用いた検索方法に関わりのない用語はどれか。 カテゴリ／キーワード／プロトコル／絞り込み
77.7	1	情報が持つ性質として、当てはまらないものはどれか。 みんなが知っている情報は価値が高い／情報は複製が可能である／情報は伝達できる／情報は人により価値が異なる
77.0	4	不特定多数に公開する Web サイトを作成する際に注意する点として、最も適切でないものはどれか。 更新年月日を明示する／作成者の個人情報を明示する／目的に合った Web サイトを作成する／定期的に更新する
76.7	18	コンピュータネットワークなどの情報技術を利用して行う学習の形態を何とというか。 電子政府／e コマース／e ラーニング／e ビリング

表 3 全体正答率の低い設問

正答率 (%)	問番号	問題文と選択肢 (「わからない」の選択肢を除く)
18.0	39	情報化の進展により普及した情報システムの例として、適切でないものはどれか。 URL／ETC／ATM／POS
20.7	36	連続的に変化する量を、離散的な量で表すことを何とというか。 デジタル化／アナログ化／バイト化／ビット化
29.3	38	検索エンジンを用いた検索条件として、適切でないものはどれか。 OR 検索／AND 検索／IF 検索／NOT 検索
34.8	16	無線通信の規格でないものはどれか。 LTE／Bluetooth／WiMAX／ADSL
34.9	30	記憶装置にはならないものはどれか。 フラッシュメモリ／バーコードリーダー／RAM／ROM
36.8	27	文字、画像、音声、映像などのデータをデジタル化し、コンピュータ上で統合的に取り扱えるようにしたものを何とというか。 マスメディア／インタラクティブ／リンク／マルチメディア
37.0	19	一次情報の収集手段として、当てはまらないものはどれか。 インタビューによる収集／アンケート調査による収集／実験データの収集／新聞や書物による収集
40.4	2	出力装置であるものはどれか。 CPU／キーボード／プリンタ／ハードディスク
41.2	26	待ち合わせ場所を案内するために、友人に地図情報を伝えたい。その際利用するものとして、最も適切でないものはどれか。 電子メール／SNS／FAX／音声電話
43.1	3	いつでも、どこからでもコンピュータや情報端末が利用でき、人間の生活を支援するような情報環境を何とというか。 ユビキタス社会／イントラネット／オフィスオートメーション／e ビジネス

LTE や ADSL, 問 30 の RAM や ROM, 問 2 の CPU などである。

3.4 正答率に関する考察

これらの結果から、各設問が受験者の日常生活において慣れ親しんでいる語句や状況か否かが、正答率に影響していると考えられる。つまり、学生の日常生活で慣れ親しんでいるスマートフォンでおこなうことの多い、SNS、動画閲覧、Web 閲覧、Web 検索などに関する語句や操作やリテラシーについては、おそらく高校で教科情報として学ぶ前からある程度は知っていたりそのような状況を体験していたりして、知識が定着しやすかったと考える。一方、

POS, RAM などの, 多くの学生にとって耳なじみのない語句などは, 高校の教科情報ではじめて学んだ学生も多く, 知識として定着しにくかったと考える。

4. 設問再考

3 節より, 正答率の高い設問および低い設問に特徴が見られることから, 正答率の低い設問 20 問と全 40 問との関係を分析し, 全 40 問の構成を再考する。下位の設問を選んだ理由は, 今後も, 全体的に知識が定着している正答率の高い設問より, 正答率の低い設問の詳細を把握していくことが重要であると考えたためである。その一方で, 半数の 20 問に減らした場合においても, これまでの全 40 問と同等程度の成績の把握が可能であるという効率も考えたためである。20 問と設定した理由は, 全体の半数を現時点での目安とし, 今後本研究を進めていく上で, その問題数を減らしたり増やしたりしながら, 設問の違いによる評価の効果と効率を再考し続けることを計画しているためである。

4.1 下位 20 問の成績と全体成績との相関分析

図 1 は, 全分析対象者の全年度の正答率の低い設問 20 問の各受験者の成績 (横軸 20 点満点) とその受験者の全 40 問の成績 (縦軸 40 点満点) との相関を示す。このグラフの相関係数は 0.930 で, 線形回帰の決定率は 0.865 であり, 下位 20 問の成績と全 40 問の成績との間には強い正の相関がある。したがって, 下位 20 問に正答できていない受験者は全 40 問においても成績が悪く, 反対に, 下位 20 問に正答できている受験者は全 40 問においても成績が良いことがわかる。

この相関分析の結果は, 各受験者の下位 20 問の成績から全 40 問の成績を推定できる可能性を示している。つまり, 全体正答率の低い 20 問の設問のみのテストを実施するだけで, それらの 20 問における各受験者の詳細を把握しつつ, 全 40 問に相当する成績を把握できる可能性を示している。

4.2 下位 20 問からの全体成績の推定

4.1 節の結果を受けて, 各受験者の下位 20 問の解答パターンからその受験者の全 40 問の成績を推定するニューラルネットワークを構築する。データセットの入力データは各

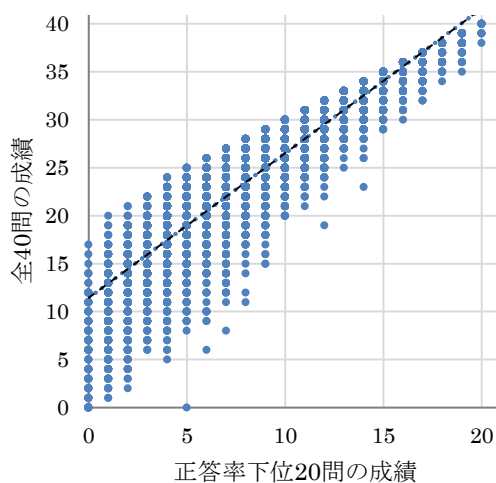


図 1 正答率下位の成績と全問の成績との関係

受験者の下位 20 問の設問の解答パターン (20 問分の解答した選択肢) である。また, 正解データは, その受験者の全 40 問の成績 (0 から 40 までの整数) である。

図 2 にこの学習に使用したベクトル分類ネットワークのレイヤの模式図を示す。Input レイヤはニューラルネットワークの入力層である。Affine レイヤは全ての入力値から全ての出力ニューロンへ結合する全結合層を示す。ReLU は入力値の正規化線形関数を出力する。Dropout は, 過学習を抑制するために, 入力の要素を指定した確率 (今回は 0.5 に設定した。) で 0 にするレイヤである。そして, SquaredError レイヤは, 全 40 問の成績として 0 点から 40 点までの連続値を推定するため, 得たい値と要素毎の二乗誤差を算出するロス関数のレイヤである。

ニューラルネットワークへの学習および評価データとして, データセットの全分析対象者のデータ (17,049 件) を年度に関係なく無作為に並べ替えた上で, そのうちの 8 割を学習用に, 残りの 2 割を評価用に使用した。ニューラルネットワークへの入力データは, データセットにおける各受験者の下位 20 問それぞれの 5 つの選択肢の内の 1 を 5 つのカテゴリの内の 1 つとし, 当てはまる選択肢を 1 つの要素とするベクトルに変換したデータである。また, 教師あり学習のための正解データは, データセットの全使用データの成績の平均と標準偏差を用いて各受験者の全 40 問の成績を標準化した値である。このネットワーク構築および学習と評価のためのソフトウェアとして, Sony Neural Network Console[8]を使用した。

図 3 に, このネットワークを使って, 評価用 3,410 件の下位 20 問の解答パターンから推定した全 40 問の成績 (横軸) と実際の全 40 問の成績 (縦軸) との関係を示す。縦軸・横軸とも, 前述の標準化した値を示している。このグラフの相関係数は 0.944, 線形回帰の決定率は 0.891 である。

4.3 成績の相関および推定に関する考察

4.1 節の下位 20 問の成績のみに基づく線形回帰の決定率より, 4.2 節の下位 20 問の解答パターンに基づくニューラルネットワークで推定した成績の決定率の方が若干高くなった。これは, 解答パターンまで使用する方が, より精度高く, 全 40 問分の成績を推定できることを示している。あ

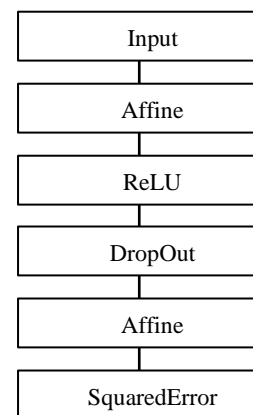


図 2 ニューラルネットワークのレイヤ

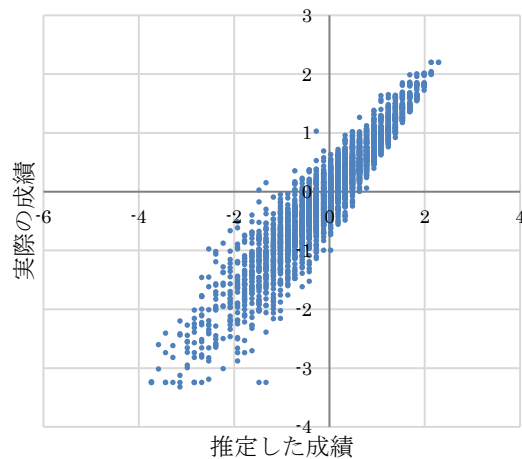


図 3 推定した成績と実際の成績

わせて、これまでの 40 問の半分の 20 問で実施したうえに、その 20 問の解答パターンから、これまでの 40 問と同等の成績を推定して把握できる可能性を示している。受験者の解答時間がおよそ半分の時間で済むことは、効率が上がることはもとより、テスト実施の簡便さもよくなると考えられるため、著者らの 4 大学のみならず、多くの大学でもこのテストを実施しやすくなることを示唆している。

5. おわりに

本論文では、4 大学 6 年間に渡る全 17,049 人分の高校普通教科情報の知識に関する五肢択一式の全 40 問のテストの解答を分析した。この結果として、正答率の低い設問には、多くの学生の日常生活では耳なじみの少ないアルファベット 3 文字程度で略された専門用語が含まれていることが多いこと、そして、全体正答率下位 20 問の解答パターンから全 40 問分の成績を決定率 0.891 で推定できることを示した。

この結果から、学習方法について、情報に関する専門用語は、ただ覚えるだけでは定着しにくいと考えられるので、正答率の高い設問のように、日常生活に関連させて学習する方法が望ましいと考える。また、設問数について、e-ラーニングシステムを使用すれば、全体正答率の低い 20 問の設問で詳細を確認し、その成績が低い受験者に対しては、残りの設問への解答を促す（反対にいうと、全体正答率の低い設問において高い成績であれば、そこでテスト終了とする）受験方法が考えられる。

正答率の低い設問の解答パターンと全 40 問の成績との間に関係があることから、今後は、設問ごとの誤答を詳細に分析して、誤りの傾向に基づく学習方法を提案したり、20 問の解答パターンから全 40 問の解答パターンを推定する方法を検討したりする。また、2022 年度から新たに共通必修科目となった「情報 I」の学習指導要領[2]の内容を精査し、プログラミング、アルゴリズム、ネットワークなどに関する知識を問う設問の更新や追加も検討する。

謝辞

本研究は JSPS 科研費 22K02831 の助成を受けたものです。

参考文献

- [1] 中野 由章, “高等学校共通教科情報科の変遷と課題”, 情報処理, Vol.59, No.10, p.933 (2018).
- [2] 文部科学省, “高等学校学習指導要領(平成 30 年告示)解説 情報編”, (2017).
- [3] 独立行政法人大学入試センター, “平成 30 年告示高等学校学習指導要領に対応した令和 7 年度大学入学共通テストからの出題教科・科目について”, (2021).
- [4] 小侯 昌樹, 吉川 雅修, 金子 大輔, 石田 雪也, 西端 律子, “8 大学共通の情報教科プレースメントテストにおける出題範囲の検証”, 日本情報科教育学会 (JAEIS) 第 7 回全国大会講演論文集, pp.95-96 (2014).
- [5] 金子 大輔, 石田 雪也, 小侯 昌樹, 吉川 雅修, 古賀 崇朗, “大学の初年次学生を対象とした情報に関する基礎知識調査の開発と調査結果の分析”, 日本教育工学会論文誌, Vol.40, pp.201-204 (2017).
- [6] 吉川 雅修, 小侯 昌樹, 石田 雪也, 西端 律子, 金子 大輔, “新入生に対する情報教科プレースメントテスト結果における学部間の傾向分析”, 日本情報科教育学会 (JAEIS) 第 7 回全国大会講演論文集, pp.93-94, (2014).
- [7] 小侯 昌樹, 吉川 雅修, “大学初年次情報授業の実施前テストと実施後テスト — 山梨大学全学実施を事例として —”, 情報教育, Vol.2, pp.41-50, (2020).
- [8] Sony, “Neural Network Console”, <https://dl.sony.com/>, (最終閲覧日: 2022 年 6 月 16 日).