

CBTにおける問題提示の一覧性に注目した比較

Comparative study of Presentation Methods for Examination Questions in CBT

加藤 尚吾*
Shogo Kato*

加藤 由樹**
Yuuki Kato**

東京女子大学現代教養学部／教育テスト研究センター*
相模女子大学学芸学部／教育テスト研究センター**
School of Arts and Sciences, Tokyo Woman's Christian University
/ Center for Research on Educational Testing (CRET)*
Faculty of Arts and Sciences, Sagami Women's University
/ Center for Research on Educational Testing (CRET)**

<あらまし> 本研究では、CBTにおける問題提示の一覧性に注目した実験を行った。具体的には、以下の3つに群分けをして試験を行い、それぞれの成績を比較した。(1)PCを1台使用し、従来のように、1つのディスプレイで問題文、設問、解答フォームの順番に提示する群、(2)PCを1台使用し、紙で問題文を提示し、PCのディスプレイに設問と解答フォームを提示する群、(3)iPad及びPCを使用し、iPadのディスプレイには問題文、PCのディスプレイには設問と解答フォームを提示する群の3つである。これらの比較の結果、試験の点数には群間で差は見られなかった。一方、試験後のアンケート結果から、紙で問題を提示されPCで解答をした群の被験者が、試験の煩わしさを感じていたことがわかった。

<キーワード> Computer Based Testing (CBT) 一覧性 問題提示 メディア比較

1. はじめに

今後、コンピュータベースの試験(CBT)がさまざまな場面で導入されると予測される。CBTでは、従来の紙と鉛筆とは異なる問題提示や解答の仕方が、コンピュータを用いることで可能になる。例えば、同じPCでも、キーボード入力もあれば、タッチペンを用いたものもある。更には、紙に書いた文字を認識できるデジタルペンもある。そのため、従来の紙と鉛筆の試験とそれらのメディアを比較し、メディアの影響を調べる研究は重要であると考えられる(加藤ほか 2010)。

そこで、加藤ほか(2010)は、試験における問題提示および解答記入のメディアとして、デジタルペン、タブレットPC、PC、紙と鉛筆の4種類を比較した。具体的には、被験者96名を、それぞれのメディアを用いて試験問題を解答する4群に人数・性別ともにほぼ均等になるように分け、群間で試験の成績等を比較した。この結果、(1)紙の問題文(文章問題の文章など)があり、別に紙の解答用紙があるもの(すなわち、紙と鉛筆を使用した群およびデジタルペンを使用した群)と、(2)PC

のディスプレイに問題文に続いて解答フォームがあるもの(タブレットPCを使用した群およびPCを使用した群)の間に、成績に差があった。その差は、問題文を参照しながら抜き出したりまとめたりすることを求めた問題で見られ、(2)で点数が低かった。この差について、問題を一覧的に見られるかどうかによる影響が考えられた。

2. 目的

本研究では、CBTにおける問題の一覧性に注目した実験を行った。具体的には、以下の3つに群分けをして試験を行い、それぞれの成績を比較した。(1)PCを1台使用し、従来のように、1つのディスプレイで問題文、設問、解答フォームの順番に提示する群、(2)PCを1台使用し、紙で問題文を提示し、PCのディスプレイに設問と解答フォームを提示する群、(3)iPad及びPCを使用し、iPadのディスプレイには問題文、PCのディスプレイには設問と解答フォームを提示する群の3つである。表1に、それぞれの群名及び試験の提示方法、問題文の一覧性の高さをまとめた。

表 1 本実験で用いた 3 群

| 群名 | 1 ディス プレイ | デ イス プレ イ紙 | 2 デ イス プレ イ |
|--------------------------|-----------------|---------------------|-------------------------|
| 提 示 方 法 | PC | 紙 | iPad |
| 問 題 文 設 題 及 び 解 答 フォ ー ム | | PC | PC |
| 問 題 文 の 一 覧 性 | 低 | 高 | 高 |

3. 方法

本実験の被験者は、リクルート会社によって集められた 60 名の大学生（男女各 30 名）であった。被験者らを、性別や専攻が均等にばらつくように配慮して、問題の提示方法の異なる 3 つの群（表 1 参照）に、20 名ずつに分けた。彼らは、メディアの使用方法等の説明を受けた後に、PISA2003（文部科学省 2003）の「読解力」試験問題の解答を行った。この試験問題は、3 種類の日本語の長文（インフルエンザ、ランニングシューズ、落書きに関する長文）に関する多肢選択問題 6 問、記述問題 7 問から構成されていた。試験終了

後、彼らはアンケートに回答した。なお、アンケート項目については、「4.3. アンケートの結果」において具体的に示す。

本研究では、被験者の解答及び試験後に行ったアンケートへの回答について、群間で比較分析した。

4. 結果

4.1. 多肢選択問題の結果

表 2 に、多肢選択問題について各群の正答率を示す。なお、表中の設問名を、例えばインフルエンザに関する長文の設問①をインフル 1、ランニングシューズに関する長文の設問①をランニング 1 と省略した（以後、図表内で設問を表す際には同様の省略を用いた）。

設問ごとに群間で正答率を比較したところ、どの設問においても有意な差は見られなかった。

4.2. 記述問題の結果

初めに、記述問題の各設問の平均文字数を、群ごとに調べた（図 1 及び表 3 参照）。

設問ごとに群間で平均文字数を比較したところ、特に、落書きの設問③で 1 ディスプレイ群と 2 ディスプレイ群の間の差に有意傾向 ($F(57, 2)=2.48, p<0.10, \text{Tukey HSD}$) が

表 2 多肢選択問題における各設問の正答率の群間での比較

| 多肢選択問題 | インフル 1 | インフル 3 | インフル 5 | ランニング 1 | ランニング 4 | 落書 1 |
|-----------|--------|--------|--------|---------|---------|--------|
| 1 ディスプレイ群 | 100.0% | 45.0% | 75.0% | 100.0% | 70.0% | 80.0% |
| ディスプレイ紙群 | 85.0% | 50.0% | 90.0% | 100.0% | 85.0% | 100.0% |
| 2 ディスプレイ群 | 100.0% | 50.0% | 80.0% | 90.0% | 85.0% | 95.0% |

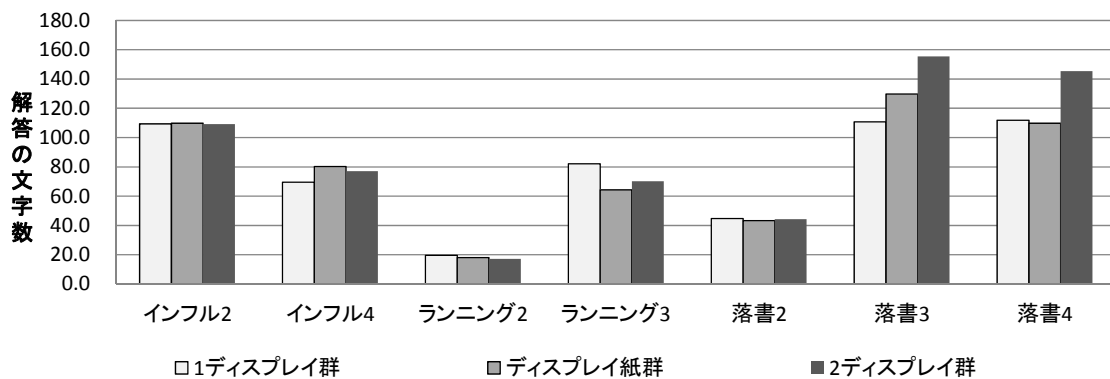


図 1 記述問題における各設題の解答文字数の平均の群間での比較

表 3 記述問題における各設題の解答文字数の平均の群間での比較

| 記述問題 | インフル 2 | インフル 4 | ランニング 2 | ランニング 3 | 落書 2 | 落書 3 | 落書 4 |
|-----------|-----------------|----------------|----------------|----------------|----------------|-----------------|-----------------|
| 1 ディスプレイ群 | 109.3 (38.9) | 69.5 (27.0) | 19.6 (20.5) | 82.0 (36.6) | 44.7 (21.7) | 110.7 (50.2) | 111.8 (38.6) |
| ディスプレイ紙群 | 109.8 (51.8) | 80.3 (44.0) | 18.1 (16.1) | 64.4 (38.2) | 43.4 (25.3) | 129.8 (68.8) | 109.8 (75.5) |
| 2 ディスプレイ群 | 109.2 (47.1) | 77.0 (42.3) | 17.1 (8.9) | 70.2 (23.4) | 44.2 (23.5) | 155.4 (66.8) | 145.4 (74.4) |

注) 下段の括弧内の数値は標準偏差である。

見られた。また、落書きの設問④でディスプレイ紙群と 2 ディスプレイ群の間に有意差 ($F(2, 57)=3.51, p<0.05, \text{Tukey HSD}$) が見られた。これらの差では、どちらにおいても 2 ディスプレイ群の文字数がより多かった。

続いて、図 2 及び表 4 に、記述問題における各群の平均点を示す。設問ごとに群間で平均点を比較したところ、どの設問においても有意な差は見られなかった。

4.3. アンケートの結果

図 3 の横軸及び表 5 の 1 列目に示された各項目について 5 段階評定 (1: あてはまらない

い~5: あてはまる) で回答を求めた。

各群で評定値を比較したところ、「実力を発揮できた」で 1 ディスプレイ群とディスプレイ紙群の間の差に有意傾向 ($F(57, 2)=2.88, p<0.10, \text{Tukey HSD}$) が見られた。「解答が記入しやすかった」で 1 ディスプレイ群とディスプレイ紙群及び 2 ディスプレイ群の間に有意差 ($F(57, 2)=5.15, p<0.05, \text{Tukey HSD}$) が見られた。また、「集中できた」で 1 ディスプレイ群とディスプレイ紙群の間の差に有意傾向 ($F(57, 2)=3.09, p<0.10, \text{Tukey HSD}$) が見られた。そして、「イライラした」で 1 ディスプレイ群とディスプレイ紙群の間に有

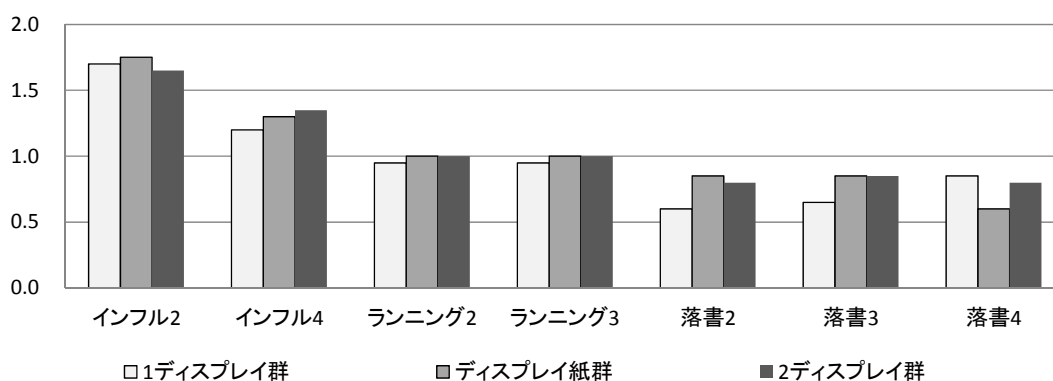


図 2 記述問題における各設問の平均点の群間での比較

注) インフル 2 及びインフル 4 のみ 2 点満点、他は 1 点満点である。

表 4 記述問題における各設問の平均点の群間での比較

| 記述問題 | インフル 2 | インフル 4 | ランニング 2 | ランニング 3 | 落書 2 | 落書 3 | 落書 4 |
|-----------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| 1 ディスプレイ群 | 1.70 (0.46) | 1.20 (0.87) | 0.95 (0.22) | 0.95 (0.22) | 0.60 (0.49) | 0.65 (0.48) | 0.85 (0.36) |
| ディスプレイ紙群 | 1.75 (0.43) | 1.30 (0.71) | 1.00 (0.00) | 1.00 (0.00) | 0.85 (0.36) | 0.85 (0.36) | 0.60 (0.49) |
| 2 ディスプレイ群 | 1.65 (0.48) | 1.35 (0.79) | 1.00 (0.00) | 1.00 (0.00) | 0.80 (0.40) | 0.85 (0.36) | 0.80 (0.40) |

注) 下段の括弧内の数値は標準偏差である。

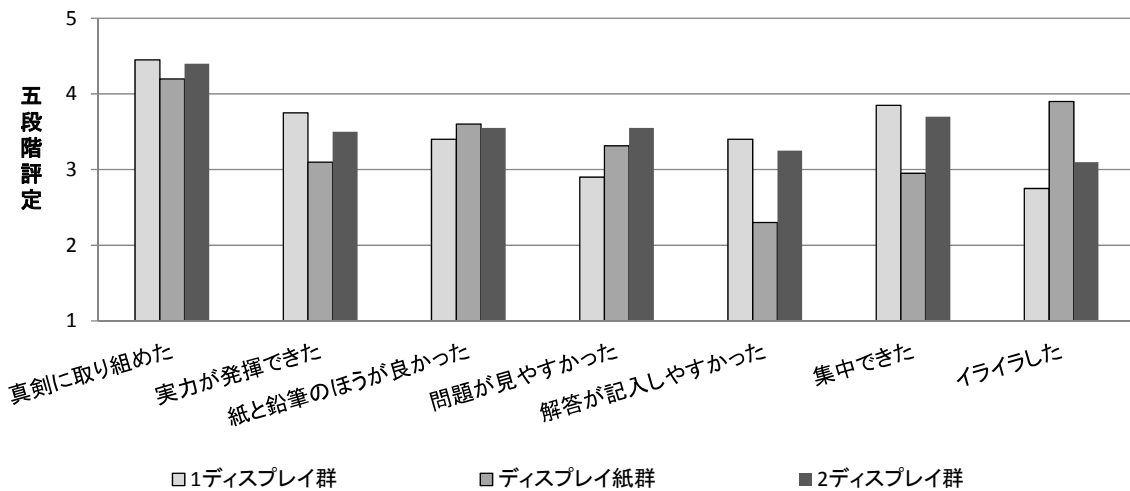


図3 アンケートにおける各項目の平均評定値の群間での比較

表5 アンケートにおける各項目の平均評定値の群間での比較

| 質問項目 | 1ディスプレイ群 | ディスプレイ紙群 | 2ディスプレイ群 |
|--------------|-------------|-------------|-------------|
| 真剣に取り組めた | 4.45 (0.74) | 4.20 (0.93) | 4.40 (0.73) |
| 実力が発揮できた | 3.75 (0.70) | 3.10 (1.04) | 3.50 (0.74) |
| 紙と鉛筆のほうが良かった | 3.40 (1.07) | 3.60 (1.02) | 3.55 (1.28) |
| 問題が見やすかった | 2.90 (1.04) | 3.32 (1.08) | 3.55 (0.97) |
| 解答が記入しやすかった | 3.40 (0.97) | 2.30 (1.00) | 3.25 (1.41) |
| 集中できた | 3.85 (1.01) | 2.95 (1.36) | 3.70 (1.19) |
| イライラした | 2.75 (1.44) | 3.90 (1.09) | 3.10 (1.30) |

注) 括弧内の数値は標準偏差である。

意差 ($F(57, 2)=3.98, p<0.05$, Tukey HSD) が見られた。

5. 考察

本研究では、CBTにおける問題の一覧性に注目した。しかし、得られた結果から、一覧性が試験の点数に影響を与えたとは言えない。

しかし、アンケート結果からは、いくつか示唆されることがあると考えられる。アンケート結果を概観すると、ディスプレイ紙群の被験者は、解答の記入をしにくく、イライラして集中できず、実力を発揮できなかつたとより感じていたと読み取ることができる。すなわち、紙で問題文を提示され解答をディスプレイで行う方法は、被験者に煩わしさを与える。しかし、そのような環境でも、被験者が耐えて試験に取り組んだ結果、点数上は他の群との間に差が見られなかったと推察される。この煩わしさの要因として、限られたスペースである机の上で、紙

をめくる動作と、画面をスクロールしキーボードを打つ動作が混在していることがあると考えられる。すなわち、紙はそれ自体を物理的に動かして使用するメディアであり、PCは設置され画面内を操作するメディアである。この点について、今後さらなる検討を重ねていきたい。

最後に、一覧性について考える。一般に、一覧性が高いと言われるメディアには紙の新聞や雑誌、書籍が挙げられる。一方、本実験で使った問題文は、上下のスクロールで全体を見られるA4縦方向・横書きのページ設定で作成されたものであった。従って、一覧性をより求められる問題について、今後探っていきたい。

参考文献

- 加藤由樹, 加藤尚吾, 赤堀侃司, 吉本真代, 杉山康彦 (2010) デジタルペン、タブレットPC、PCおよび紙と鉛筆の4種類のメディアを用いた試験に関する比較分析. 日本教育工学会第26回全国大会講演論文集, pp.681-682
 文部科学省 (2003) PISA (OECD生徒の学習到達度調査), 東京

