

能力検査受検にスマートフォンを使用することの影響とその要因

○園田 友樹, 藤田 彩子

株式会社リクルートキャリア測定技術研究所

1. 研究の背景と目的

総務省の平成26年通信利用動向調査(2015)によれば, 20代のインターネット利用の87.5%はスマートフォン(以下SMP)によるものが占めている. 株式会社リクルートキャリアの調査(2016)によれば, 大学生の94.8%がSMPを保有しており, 就職活動をする学生にとってもSMPがメインデバイスとなっている. 現在, Web上で受検するアセスメントやテストは採用場面において広く利用されている. SMPやタブレットなどの普及に伴い, 受検にPCではなくSMPやタブレットが用いられる場合も増えると考えられる.

一方で, SMPでの受検, とりわけ能力検査の受検においていくつかの課題も指摘されている. Kingら(2014)の研究によると, Situational Judgmentなど, テストによっては, PC受検とモバイルツール受検の結果が同等とみなせるが, 能力検査についてはモバイルツールが不利となる. LaPort(2016)の研究によれば, 能力検査では問題の画面表示のされ方が得点に影響する事が確認されている. また, 採用場面での利用を想定した時, 受検者側の納得感も重要である. 受検環境によっては受検者の納得感を損ね, 不公平感を抱かせる可能性がある.

SMP などモバイルツールによるアセスメント受検については, 海外を中心にここ数年実証的な研究が行われているが, 国内での研究はほとんど見当たらない. 本研究では, 国内の能力検査について, PCで受検した場合とSMPで受検した場合でどのような違いが生じるか検証した. 具体的には, デバイスの違いが問題の難易度にもたらす影響とその要因, および受検者からみた回答しやすさ・しにくさの程度とその要因について検証した.

2. 手続き

2-1. 調査項目

2-1-1.能力検査 本研究では, (株)リクルートキャリア(旧人事測定研究所)が開発した, 基礎能力検査 GAT (General Ability Test)の Web テスト 1 版を用いた. これは, PC での受検を想定して開発された Web 配信のテストである. GAT は, 言語尺度と非言語尺度の二つの下位尺度で構成されている. 言語については, 「語彙」「語句の並べ替え」「文章の論理展開や論旨の把握」などの下位領域で構成される. 非言語については, 「定価・速度の計算などの数量的情報の取り扱い」「表の計算」「確率・組合せ」「論理・推論」などの下位領域で構成される.

2-1-2. 「回答のしやすさ」に関するアンケート 能力検査回答後, テストの回答のしやすさについて問うために「お使いの端末での回答のしやすさはいかがでしたか?」という質問に対し, 「1:とてもやりやすかった」から「5:とてもやりにくかった」の5段階で回答してもらった. 「4:やややりにくかった」「5:とてもやりにくかった」と回答した対象者からは, 回答しにくさの具体的内容について自由記述で回答を得た.

2-2.調査の設計および対象者

今回は回答負荷を考慮し、言語検査(36問, 制限時間 30分)を回答する群と非言語検査(22問, 制限時間 45分)を回答する群の2つに対象者を分けてインターネット調査により回答を得た。実施形態は、①PC、②PC画面をそのまま表示させたSMP(以下SMP[PC])、③SMP用に調整した画面を表示させたSMP(以下SMP[調整])の3つで、対象者にはランダムに①～③の実施形態を割り振った。同一人物に複数の実施形態での回答を依頼しなかったのは、同じ内容の問題を複数回受検することによる練習効果を排除する為である。

2-3. データのスクリーニングとサンプリング

まず、得られた合計2,799名のデータから、ほとんど同じ回答を入力している等の明らかに回答傾向の偏った196名を除外した。実施形態をランダムに割り振ったが、PC群とSMP[PC]・SMP[調整]群との間に、性別・年齢・学歴について割合が多少異なっていた。分散分析の結果、言語検査で学歴に、非言語検査では性別・年齢・学歴のすべてについて得点に有意な差が見られた為、PC群についてランダムサンプリングを行い、PC群の各属性の割合がSMP[PC]・SMP[調整]群の割合と同程度になるよう調整した。その結果、実際に分析に用いたデータの①～③の各群の属性ごと人数の内訳は【図表1】のようになった。

3. 結果と考察

3-1. 項目特性値に及ぼす影響

言語尺度、非言語尺度それぞれにつき、項目特性値(特に困難度パラメータb)の比較を行った。項目特性値の算出には、Rのパッケージ(difR)を使用した。

困難度パラメータbの値を比較すると、言語の問題、非言語の問題ともに、平均的にはPCで受検した場合が最も低く、次いでSMP[PC]、SMP[調整]となった(【図表2】参照)。問題が同じでも実施形態によって困難度が変化するといえる。その程度は平均すると困難度パラメータの0.3前後の違いで、ハイステークスなテストの場合は、無視できない違いと思われる。個々の問題を見ていくと、言語・非言語ともに、問題の低位領域による特徴的な傾向は見られず、各問題で異なる傾向が見られた。【図表3】は、横軸がPC受検での困難度、縦軸にSMP受検での困難度(●はSMP[PC]、○はSMP[調整])をプロットしたものである。PCとSMPで困難度の大きな差があった問題数題(【図表3】「非言語」のグラフ中でA、B、Cの項目)について、困難度差が生じた要因について考察した。

《項目A》SMP[調整]はPCと同程度の困難度だが、SMP[PC]では困難度がかなり上がっている。問題画面を確認したところ、「-(マイナス)」の記号がSMP[PC]で見づらく、これを見落としたことによる不正解が多くなったことが考えられる(実際、マイナスがない場合に導かれる値を答えている受検者が他と比べて非常に多かった)。

《項目B》SMP[PC]はPCよりやや困難度が低いが、SMP[調整]

【図表1】各テストの人数集計

●言語検査

【性別】

	男性	女性
PC	118	275
SMP[PC]	118	275
SMP[調整]	106	247

【年齢】

	22～24歳	25～29歳	30～34歳
PC	30	148	215
SMP[PC]	38	153	202
SMP[調整]	34	130	189

【学歴】

	大卒	大学院卒
PC	354	39
SMP[PC]	346	47
SMP[調整]	323	30

●非言語検査

【性別】

	男性	女性
PC	93	218
SMP[PC]	117	281
SMP[調整]	122	276

【年齢】

	22～24歳	25～29歳	30～34歳
PC	19	119	173
SMP[PC]	34	155	209
SMP[調整]	36	145	217

【学歴】

	大卒	大学院卒
PC	264	47
SMP[PC]	348	50
SMP[調整]	357	41

【図表2】実施形態ごと項目特性値の平均値

		識別力	困難度
言語尺度	PC	1.23	-0.91
	SMP[PC]	1.12	-0.73
	SMP[調整]	1.24	-0.66
非言語尺度	PC	1.65	0.22
	SMP[PC]	1.60	0.35
	SMP[調整]	1.70	0.58

では困難度が高くなっている。この問題は、表を見て問いに答える2問組の問題の2問目で、SMP[調整]では表と問題を同時に見ることができず、スクロールしながら解かなければならない。

《項目C》SMP[PC]はPCとはほぼ等しい困難度だが、SMP[調整]では困難度が高くなっている。この問題は、分数を答えさせる問題で、PCとSMP[PC]は、回答欄の□/□が1行に収まっているが、SMP[調整]では「/」の後で改行され、一見分かりにくい回答欄となっていた。

単にSMPに適応した画面にすることが、必ずしも問題の難易度の低下にはつながっていないことが示唆される。A、B、C以外の問題についても、SMP[調整]でSMP[PC]より困難度が高くなるものが多く見られ、改行位置やフォントなど、細かいレイアウトの影響は思った以上に大きいかもしれない。

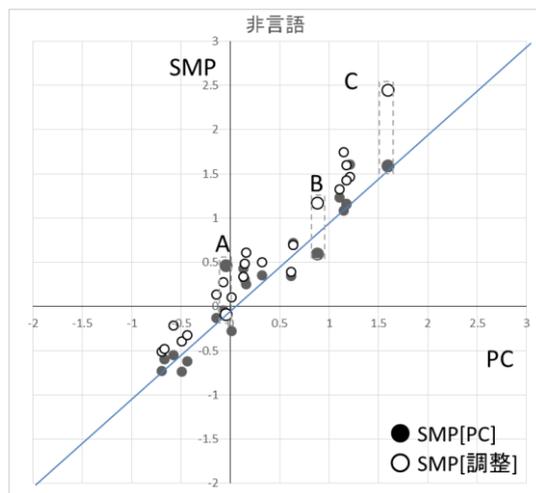
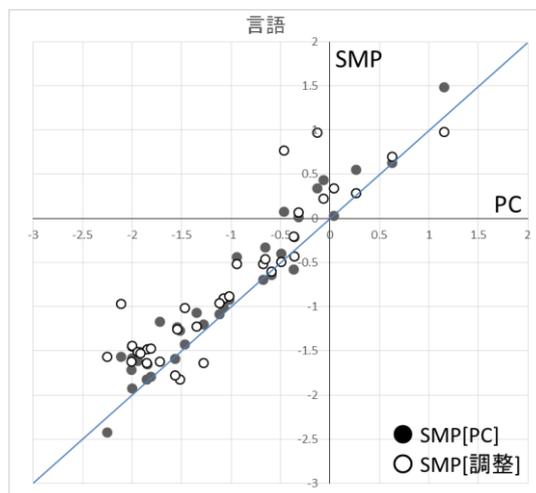
3-2. 回答のしやすさに及ぼす影響

「回答のしやすさ」について実施形態による違いがあるか、言語・非言語それぞれについて分散分析を行った。その結果、言語($F(2,1136)=11.794, p<.001$)、非言語($F(2,1104)=16.754, p<.001$)共に実施形態による有意差があった。Tukeyを用いた多重比較によれば、言語、非言語のいずれのテストについても回答のしやすさは①PCと③SMP[調整]は同程度で、②SMP[PC]が有意に回答しにくい結果となった。

回答しにくさの具体的内容を捉えるために、自由記述の回答の形態素解析を行った。形態素解析には RMeCab を用いた。全形態素の中から名詞・動詞・形容詞に該当するものについて出現頻度を計量し、上位10までの形態素を確認した。結果は【図表4】および【図表5】の通りであった。言語問題では、PCでの回答については問題そのものの文字量や読みにくさについて指摘するもの(「長い文章問題が読みにくい」等)が中心であった。一方で、SMP[PC]では文字の小ささに加え、画面サイズ(「画面が小さくイライラする」等)や選択肢ボタン(「ボタンが小さくて打ちづらい」等)についての言及が特徴的であった。【図表4】からも②SMP[PC]に特徴的な言葉として、「拡大」「ボタン」が挙げられる。SMP[調整]についても同様に画面が小さいことによる文字の読みにくさについてのコメントが多く見られた。【図表4】において、②SMP[PC]、③SMP[調整]にはあるが①PCにはない言葉として、「画面」「見にくい」「字」があり、また、トップに共通して「小さい」があるように、画面をSMP用にしても、やはり見づらさを感じさせていることがうかがえる。

非言語問題では、PCでの回答については、問題の難しさや計算の煩わしさについてのコメントが多い(【図表5】においても、①PC特有の言葉として「難しい」「わかる」「時間」などがある)。一方、SMP[PC]での回答では「拡大したりスクロールしたりという作業が手間」といった画面操作に関するものや文字入力についてのコメントが多く

【図表3】PC/SMP実施の各項目の困難度の違い



見られた。SMP[調整]についても文字の小ささ、文字入力に関するコメントが見られている。【図表 5】では、②SMP[PC]、③SMP[調整]だけに共通する言葉として、「画面」「小さい」「文字」「入力」「しまう」「なる」などがあり、やはり画面が小さい、入力がしにくい、という印象を持つ受検者が多いと思われる。また、SMP[PC]、SMP[調整]に共通して、「計算しているうちに画面が消えた」「すぐにロックがかかる」といった、問題を解く間の画面制御に関するコメントが見られた。

アンケート回答の結果からは、PC と SMP[調整]において、回答しやすさに有意な差がなかったことから、SMP での受検が必ずしも回答しにくい印象を与えるわけではないが、SMP ならではの回答に与える影響が明らかになった。SMP の画面上での文字認識・回答操作の困難さ、問題文面の量に起因するスクロール等の操作を要すること回答への負荷を高めることが想定され、画面上での問題表示には十分な配慮が必要である。また、一定時間を経過すると画面のロックがかかるなどの問題もあり、開始前のインストラクションなど含めてきめ細やかな対応が必要である。加えて、実施形態の違いは項目特性値ひいてはテスト得点に、ある程度の影響を与えると考えられ、実施形態として PC と SMP の両方を想定する場合は、実施形態の影響の大きい項目の発見・排除、また、実施形態の影響が小さい問題の作成などが求められる。

4. 今後に向けて

本研究では、SMP での能力検査受検における課題の一部が明らかになったが、どのような問題項目で実施形態による影響が大きくなるか等については、はっきりしないことも多かった。今後に向けては、実施形態に影響を受けやすい問題の特徴を明らかにしていくとともに、SMP での受検に適した画面設計やインストラクションの在り方、回線や画面制御など想定される他の影響についても検証して行きたい。

5. 参考文献

- Kate LaPort (2016) Mobile Assessment Comparing Traditional Cognitive, Cognitive Reasoning, and Non-Cognitive Performance. Mobile Equivalence: Expanding Research across Assessment Methods, Levels and Devices. Symposium presented at the 31th Annual Conference of the Society for Industrial and Organizational Psychology, Anaheim, CA.
- 株式会社リクルートキャリア (2016) 大学生の実態調査 2016
- King, D. D., Ryan A. M., Kantrowitz, T., & Grelle, D. (2014). MIT Versus PCIT: Assessing equivalence, individual differences, and reactions. In T. Kantrowitz & C. M. Reddock (Chairs), Shaping the future of mobile assessment: Research and practice update. Symposium presented at the 29th Annual Conference of the Society for Industrial and Organizational Psychology, Honolulu, HI.
- 総務省 (2015) 平成 26 年通信利用動向調査
- 豊田秀樹 (2013) 『項目反応理論[中級編]』p78-p89 朝倉書店
- 日本テスト学会第 14 回大会発表論文抄録集 P164-167

【図表 4】言語検査についての形態素解析結果

①PC		②SMP [PC]		③SMP [調整]	
形態素	順位	形態素	順位	形態素	順位
読む	1	小さい	1	小さい	1
文字	1	する	2	文字	2
問題	3	文字	3	する	3
文章	4	読む	4	問題	4
にくい	5	拡大	4	読む	5
小さい	6	にくい	6	画面	5
する	7	画面	6	にくい	7
づらい	8	見にくい	8	文章	8
いる	9	ところ	8	ある	9
多い	10	ボタン	8	見にくい	10
ある	10	字	8	字	10

【図表 5】非言語検査についての形態素解析結果

①PC		②SMP [PC]		③SMP [調整]	
形態素	順位	形態素	順位	形態素	順位
問題	1	する	1	画面	1
計算	2	小さい	2	する	1
する	3	画面	3	入力	3
難しい	4	文字	4	小さい	4
時間	5	入力	5	文字	5
わかる	6	にくい	6	数字	6
ところ	7	なる	7	問題	6
にくい	8	問題	8	なる	6
ない	9	計算	9	にくい	9
づらい	10	しまう	9	ところ	10
				しまう	10