

# 高大接続を視野に入れたタブレットを用いる評価問題の試作 (3)

——ペーパーテストとの比較——

安野 史子 (国立教育政策研究所), 山下 卓弥 (富山県立富山中部高等学校),  
柳澤 秀樹 (駒場東邦中学校高等学校), 高木 繁 (名古屋工業大学), 中島 範行 (富山県立大学),  
林 誠一 (富山県立砺波高等学校), 松原 静郎 (桐蔭横浜大学)

本研究は、CBTの特性を生かした教科・科目ベースの問題を試作し、検討することを目的として実施してきている。今回は、映像やシミュレーション実行のための動的オブジェクトを含む化学の問題について、PBTと評価がどのように異なるのかを探るため、CBT用に開発した問題の一部をPBT版に直し、比較調査を実施した。その結果、冊子全体としては明確な違いは確認できなかったが、PBTとCBTで問題の提示方法が異なることもあり、個別の設問において差異が観察された。

## 1 研究の目的

本研究は、昨年(2016年)度に引き続き、CBT (Computer-Based Testing)の特性を生かした教科・科目ベースの問題を試作し、検討することを目的として実施された。今回は、化学における同一内容の問題で、CBTとPBT (Paper-Based Testing) では測定結果にどのような違いがあるかを検討することを主目的とする。具体的には、高等学校の化学の履修を終えている大学1,2年生を対象とした調査を通じて、難易度の差や評価内容に違いがあるのかを検討し、それらを規定する要因が何であるのかという問いに対する解答を探る。さらに、昨年度までは、高等学校第2学年及び第3学年の生徒を対象に調査を行っていたが、今回は、高大接続の観点から、調査対象集団を大学1,2年生とし、CBTの問題について、高校生と大学生といった異なる集団で調査した場合、難易度に違いがあるのかを探ることも目的としている。

## 2 調査のデザイン

このような目的の実現のために、同一内容の問題をCBT形式とPBT形式の両方の形式で作成し、能力同等と考えられる二つの集団に解答させることと、一受験者に、CBT形式とPBT形式の両方を解答させることが必要になる。

その二つの課題の遂行のために、昨年度までに開発したCBTの問題(安野, 2018)から選ばれた6題がPBT形式に書き換えられ、それらを3題ずつ、内容と難易度がほぼ均等になる2組に分けられた。一方、受験者は、2大学の理系学部にも所属する大学1年生及び2年生

からなるサンプルを無作為に2群に分け、各群に対しそれぞれ2組の問題の一方が実施された。

## 3 方法

### 3.1 調査対象者

高等学校の学習内容を履修し終え、入学試験で理科が試験科目となっている理系学部の大学1,2年生を調査対象とした。ただし、調査実施が大学入学試験受験後多少時間が経過していることから、大学においても化学系の科目を履修しているものとし、国公立大学2大学において実施した。

### 3.2 調査問題

昨年度までに開発したCBTの化学問題(安野, 2018)は、CBTの特性を活かした問題として開発を行った。そのため、化学実験を中心とした映像やシミュレーションを実行できる動的なオブジェクトを含んでおり、PBTでは表現できない(あるいはしにくい)問題である。特に、化学変化の様子を映像で示しているものは、ペーパーテストではその現象を問題文で明示することになるため、本質的には異なる。その中で、同一内容で比較的ペーパーテストに書き換え可能な問題を、今までに開発した15題(安野, 2018)のうちから選び出し、6題をPBT形式に改作した。それらの問題の内容は、表1の通りである。前述のように、CBT形式の「映像」を見て、変化の様子やどのような現象が起こったかを記述させる設問は、PBT形式では出題が難しいため、問題ID01～03は、PBT形式への書き換えにより小問数が減少した(表2)。また、問題ID04ではPBT形式への書き換えにより設問の順番が入れ替わった(表2)。例えば、

図1は問題ID03のCBT形式の問題，図2はそれをPBT形式に書き換えた問題を示したものである。CBT形式では映像を見て，沈殿が生じている操作(例えば，図3)をすべて選ぶが，PBT形式ではそのような設問は出題できず，問題文に各操作で沈殿が生じるか生じないかを明記している。さらに，問題ID05はCBT形式では動的オブジェクトを用いたシミュレーション問題であり，インタラクティブにデータを収集すると，自動的にグラフにプロットされるが，PBT形式では同一の表現が困難なため，実験結果データの表をあらかじめ明示し，それをもとにグラフに描かせる問題としている(表2)。

これらの問題を内容と難易度がほぼ均等になる

**第3問**

$Ag^+$ 、 $Al^{3+}$ 、 $Ba^{2+}$ 、 $Cu^{2+}$ 、 $Fe^{3+}$ のうちのいくつかを含んでいる混合溶液がある。各操作の映像を見て，以下の問い(1)~(3)に答えよ。なお，前処理が2か所省略されている。また，それぞれの操作の後には，沈殿の有無にかかわらず過し，そのろ液を用いている。

(1) 操作1~5の中で沈殿が生じたのはどの操作か。すべて選びその番号を答えよ。

(2) 混合溶液に含まれている金属イオンをすべて答えよ。

(3) 前処理(ア)(イ)をどの操作の直前に行う必要があるか，その操作の番号をそれぞれ答えなさい。また，その理由を説明せよ。  
 (ア) 水溶液を煮沸し，希硝酸を加える。  
 (イ) 水溶液を酸性にする。

混合溶液

↓

操作1 塩酸を加える

↓

沈殿 ろ液

↓

操作2 硫化水素を加える

↓

沈殿 ろ液

↓

操作3 水酸化ナトリウム水溶液を加える

↓

沈殿 ろ液

↓

操作4 アンモニア水を加える

↓


沈殿 ろ液


↓


操作5 硫化アンモニウム水溶液を加える


↓


沈殿 ろ液

 (25秒)

 (65秒)

 (39秒)

 (45秒)

 (50秒)

A冊子 第3問 (終り)

図1 CBT 問題例 (問題ID03)

**第6問**

$Ag^+$ 、 $Al^{3+}$ 、 $Ba^{2+}$ 、 $Cu^{2+}$ 、 $Fe^{3+}$ のうちのいくつかを含んでいる混合溶液に次の操作1~5を行った。以下の問い(1),(2)に答えよ。

操作1 混合溶液に塩酸を加えたところ，白色の沈殿が生じた。

操作2 操作1のろ液に硫化水素を加えたところ，黒色の沈殿が生じた。

操作3 操作2のろ液に水酸化ナトリウム水溶液を加えたところ，赤褐色の沈殿が生じた。

操作4 操作3のろ液にアンモニア水を加えたところ，沈殿は生じなかった。

操作5 操作4後の溶液に硫化アンモニウム水溶液を加えたところ，沈殿は生じなかった。

混合溶液

↓

操作1 塩酸を加える

↓

沈殿 ろ液

↓

操作2 硫化水素を加える

↓

沈殿 ろ液

↓

操作3 水酸化ナトリウム水溶液を加える

↓

沈殿 ろ液

↓

操作4 アンモニア水を加える

↓

溶液

↓

操作5 硫化アンモニウム水溶液を加える

↓

溶液

(1) 混合溶液に含まれている金属イオンをすべて答えよ。

(2) 前処理(ア)と(イ)をどの操作の直前に行う必要があるか，その操作の番号をそれぞれ答えよ。また，その理由を説明せよ。  
 (ア) 水溶液を煮沸し，希硝酸を加える。  
 (イ) 水溶液を酸性にする。

図2 PBT 問題例 (問題ID03)

2組に分け，2冊子(A冊子，B冊子)を作成した。各冊子には，CBT形式とPBT形式の問題を3題ずつ割り振り，第1問から第3問はPBT形式，第4問から第6問はCBT形式の問題とし，解答時間は60分に設定した。

### 3.3 質問紙

以下の1~3の内容の質問紙も併せて作成した。ただし，質問項目はすべて選択肢式の項目である。

1. 問題について
  - 難易度，意味の理解，解答時間
2. 問題のオブジェクトについて
  - 映像時間，映像の見易さ
3. 高等学校での化学実験の頻度について

### 3.4 調査実施

- ① 調査日：2017年12月19日 (U<sub>1</sub>大学) 及び2017年12月25日 (U<sub>2</sub>大学)
- ② 調査人数：U<sub>1</sub>大学A冊子37名 B冊子38名 (1年生1学科，2年生1学科) U<sub>2</sub>大学A冊子37名 B冊子39名 (1年生2学科)
- ③ 調査方法：問題冊子(iBooksによる電子冊子及び紙媒体の冊子)，解答用紙(紙媒体)，7.9インチのタブレット端末(Apple Inc.のiPad mini2, OS:iOS9)，質問紙(マークシート)を配付し，学生の所属大学の講義室で実施した。質問紙調査は，問題を解答後に行い，調査終了後全ての配付物を回収した。

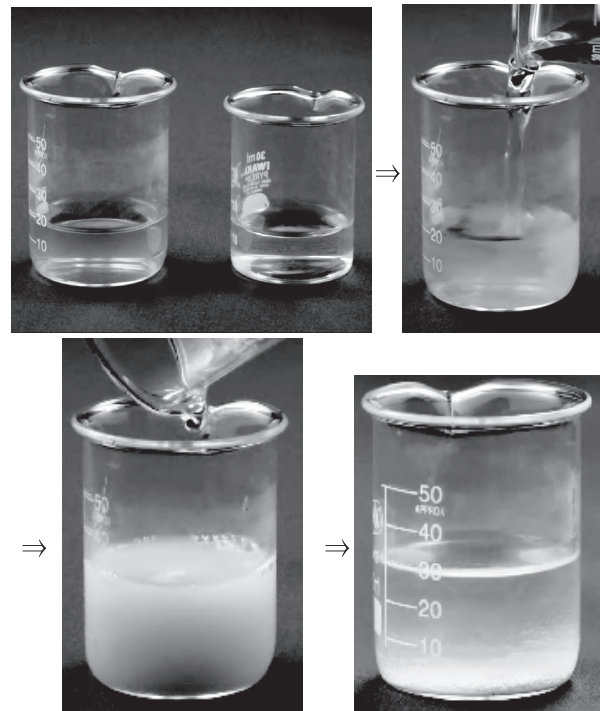


図3 CBT 映像例 (問題ID03：操作1)

表1 問題冊子の構成

問題群	問題ID	問題番号		形式		内容	CBTに含まれるオブジェクト
		A冊子	B冊子	A冊子	B冊子		
I	01	第1問	第4問	CBT	PBT	蒸気圧と沸騰(水の沸騰)	映像
	02	第2問	第5問			電池のしくみ	映像
	03	第3問	第6問			金属イオンの系統分離	映像
II	04	第4問	第1問	PBT	CBT	酢酸水溶液の中和滴定(器具・操作)	映像
	05	第5問	第2問			化学反応と量的関係	シミュレーション
	06	第6問	第3問			沈殿反応	映像

表2 調査結果(問題別)

問題ID	問題番号(配点)		小問番号	枝点	平均得点		平均得点率		平均得点 <sup>a</sup> (平均得点率)	
	A冊子	B冊子			A冊子	B冊子	A冊子	B冊子	A冊子	B冊子
01	CBT	PBT	第1問 第4問 (18) (12)	(1) 現象	2	1.57	—	78.4%	—	2.88 2.83 (24.0%) (23.6%)
				(1) 理由	2	0.76	0.81	37.8%	40.3%	
				(2) 現象	2	1.54	—	77.0%	—	
				(2) 理由	2	0.19	0.45	9.5%	22.7%	
				(3) 現象	2	1.65	—	82.4%	—	
				(3) 理由	2	0.86	0.95	43.2%	47.4%	
				(4)	2	0.65	0.34	32.4%	16.9%	
				(5) 記号	2	0.24	0.18	12.2%	9.1%	
02	CBT	PBT	第2問 第5問 (23) (17)	(1) 様子	3	2.57	—	85.6%	—	6.36 6.19 (37.4%) (36.4%)
				(1) 理由	4	1.16	0.86	29.1%	21.4%	
				(2) 様子	3	1.77	—	59.0%	—	
				(2) 反応式	3	1.72	1.58	57.2%	52.8%	
				(3)	4	1.95	1.71	48.6%	42.9%	
				(4) 正極	3	0.32	0.64	10.8%	21.2%	
				(4) 負極	3	1.22	1.40	40.5%	46.8%	
				03	CBT	PBT	第3問 第6問 (22) (17)	(1)	5	
(2)	5	3.08	3.48					61.6%	69.6%	
(3) ア	6	1.30	1.40					21.6%	23.4%	
(3) イ	6	0.08	0.00					1.4%	0.0%	
04	PBT	CBT	第4問 第1問 (24) (24)					(1)	4	1.95
				A(3)B(2) 記号	4	3.30	2.96	82.4%	74.0%	
				A(3)B(2) 値	4	2.00	2.13	50.0%	53.2%	
				A(2)B(3)	4	2.76	3.22	68.9%	80.5%	
				(4)	8	4.16	4.88	52.0%	61.0%	
05	PBT	CBT	第5問 第2問 (24) (20)	グラフ	4	3.00	—	75.0%	—	8.22 7.91 (41.1%) (39.5%)
				計算	20	8.22	7.91	41.1%	39.5%	
06	PBT	CBT	第6問 第3問 (17) (17)	(1) 組成式	4	1.68	2.49	41.9%	62.3%	6.04 8.57 (35.5%) (50.4%)
				(1) 理由	6	2.43	2.88	40.5%	48.1%	
				(2) 番号	3	0.93	1.52	31.1%	50.6%	
				(2) 反応式	4	1.00	1.71	25.0%	42.9%	

<sup>a</sup>CBT及びPBTの両方に共通している設問のみで算出

#### 4 調査結果

各問題の配点(枝点を含む)は2017年3月調査(安野,2018)の配点を踏襲した。そのため、満点がA冊子は128点、B冊子は107点となるが、本稿での比較においては、両方の冊子に共通している設問のみで行う。表3は、冊子ごとに見た形式別の得点率に関する統計量である。2組に分けた問題は、3題のまとまりの問題群で見ると、形式による得点率に差異はほとんど見られず、むしろ問題群 I (ID01 ~ 03)が問題群 II (ID04 ~ 06)よりも難度が高かったことがわかる。

次に、図4及び図5は、各冊子の問題群 I (ID01 ~ 03)と問題群 II (ID04 ~ 06)の得点率の散布図である。相関係数は、A冊子が0.54、B冊子が0.46であり、これらからも、形式による明白な違いは確認できず、問題群の難度の違いが観察できた。

さらに、表2は大問別あるいは小問別にみた結果である。大問で差が最も大きかった問題は、問題ID06(沈殿反応)で、得点率の差が14.9%ポイントで、CBT形式の方が易しいという結果であった。この問題は、CBT形式では30映像が用意されていて、知識がなくても、すべての映像を見れば正解が得られるが、PBT形式は知識がないと正解できないということに起因していると考えられる。CBT形式では観察・実験の技能、思考・判断・表現、知識・理解、関心・意欲・態度、に関して総合的に評価する問題であったのが、PBT形式では知識・理解のみの評価問題となってしまう、問題の改変により、評価していることが変質してしまった設問と言える。

小問レベルで見ていくと、問題ID06を除いた問題で差が大きかった設問は、問題ID04(1)と問題ID01(4)で、前者はガラス器具の名称(ホールピペット)を尋ねている設問で、CBT形式の得点率の方が18.9%ポイント高いという結果であった。それは、PBT形式では器具がイラストで提示されているが、CBT形式では実際に使用している様子を映像で見ることができ、器具の大きさも把握しやすいことに起因していると考えられる。形式によって、特定の誤答に偏ることなく、どちらの形式も同じように、ピペット、駒込ピペット、ビュレットといった誤答が目立ち、これらの誤答の反応率はPBT形式の方が高い。後者は、蒸気圧と沸騰の問題の(4)で、減圧後の水の沸騰に関する問いである。この問題(ID01)は、CBT形式のみ(1)~(3)において、どのような変化が起こったか等、現象を記述させているので、枝問レベルでは3問多いが、図6からもわかるように、(1)~(3)はPBT形式の方が平均得点率が高く、(4)で逆転するとともに、差が一番開き、CBT形式の方が15.5%ポイント高いという結果であった。勢いよくフラスコ内に水が引き込まれ、その水が沸騰する様子を

表3 調査結果 (形式別得点率)

問題群		I	II	全体
問題ID		01~03	04~06	
A 冊子	形式	CBT	PBT	
	平均	29.8	46.6	39.4
	標準偏差	13.3	25.1	18.0
	最低	0.0	0.0	5.6
	最高	69.6	93.4	73.8
	人数	74	74	74
B 冊子	形式	PBT	CBT	
	平均	30.2	53.1	43.3
	標準偏差	15.7	23.3	17.4
	最低	0.0	4.9	5.6
	最高	60.9	100.0	81.3
	人数	77	77	77

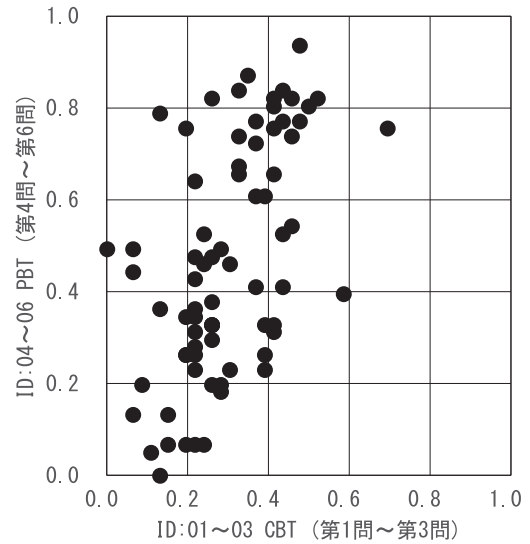


図4 問題群I (CBT)と問題群II (PBT) の得点率 (A 冊子)

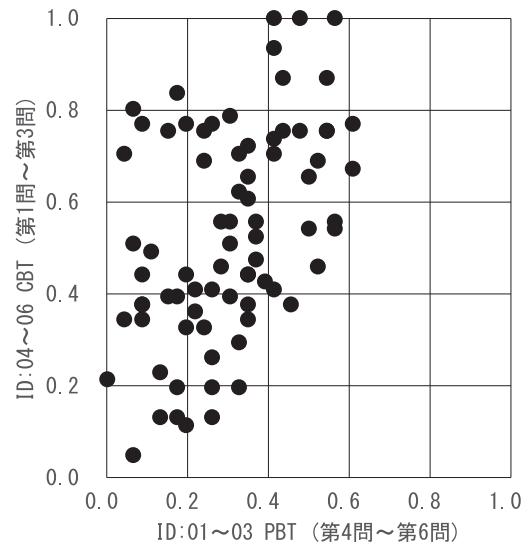


図5 問題群I (PBT)と問題群II (CBT) の得点率 (B 冊子)

表4 質問紙調査結果 (難易度)

問題群	問題ID	形式		反応平均		得点との相関	
		A冊子	B冊子	A冊子	B冊子	A冊子	B冊子
I	01			2.7	2.3	0.07	0.22
	02	CBT	PBT	2.4	2.2	0.20	0.41
	03			2.6	2.4	0.31	0.17
II	04			3.0	3.1	0.28	0.31
	05	PBT	CBT	2.7	2.3	0.42	0.54
	06			2.2	2.1	0.20	0.34

実際に見ることによって、減圧によるものであることがイメージしやすかったためではないかと考えられる。

次に、2017年3月に高校生を対象に調査を行った結果との違いについて述べる。高校生の調査は全15題で各冊子5題ずつの3冊子に分けられ調査が実施された。高校生調査においても、無作為に分けられた3組の集団に対して実施されたが、今回の調査で用いた6題は、問題群 I (ID01 ~ 03) 及び問題群 II (ID04 ~ 06) とも、高校生の調査の3冊子から1題ずつ選ばれている。各冊子の実施人数も若干名異なるため、厳密に平均得点(率)が算出できないが、各問題の平均得点を加算したもので違いを見ることにする。

問題群 I (ID01 ~ 03) (配点合計63点)は、

高校生19.2点(30.5%) , 大学生26.9点(42.7%)

問題群 II (ID04 ~ 06) (配点合計61点)は、

高校生19.5点(32.0%) , 大学生32.4点(53.1%)

で、大学生の方が、どちらの問題群も平均得点率が10%ポイント以上高く、開発したCBT問題の弁別力がある程度確認できた。

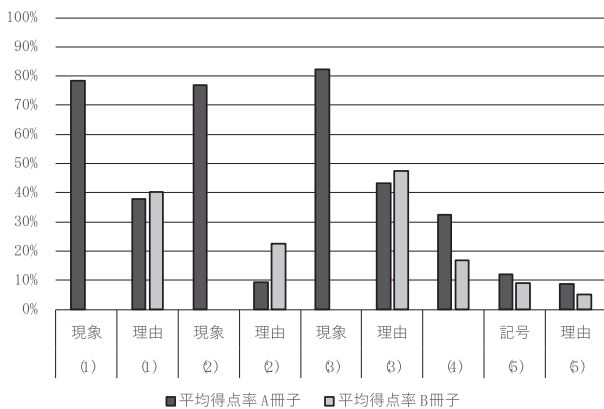


図6 小問別平均得点率 (ID01)

## 5 質問紙調査との関係

表4は、「各問題の難易度について、あなたがどのように感じたか」(難易度)を5択(1.難しい 2.やや難しい 3.ちょうどよい 4.やや易しい 5.易しい)で尋ねた質問項目への回答平均と大問の得点率との相関を示したものである。

反応平均は、2.1 ~ 3.1という結果で、「やや難しい」から「ちょうどよい」に位置しており、形式による難易度の感じ方の違いはほとんど観察できなかった。

得点との相関を、質問反応と大問得点率との相関係数でみると、問題ID05はPBT (A冊子)で0.42, CBT (B冊子)で0.54と、形式によらず高かった。この問題は、

動的オブジェクトでの実験シミュレーションによる(CBT)か問題に実験結果のデータが与えられていて、その結果を自分でプロットする(PBT)かの違いで、その後は化学計算を伴う問題である。形式による平均得点率の差は観察できなかった(表2)が、CBT形式の方が、易しいと回答している受験者ほど得点が高い傾向がより強い。

次に、形式による相関係数の差が大きかった問題は、問題ID02 (電池のしくみ)で、CBT形式(A冊子)が相関係数0.20, PBT形式(B冊子)が相関係数0.41であった。この問題も大問の平均得点率に差がほとんど見られない(表2)が、PBT形式の方が、易しいと回答している受験者ほど得点が高い傾向がより強い。図7及び図8は、横軸に質問項目の選択肢番号、縦軸に当該問題の選択肢別平均得点率、バブルの大きさに質問項目(難易度)の選択肢(5択)への回答人数をとったこの問題のバブルチャートであるが、これらの図からも、違いが明確である。「難しい」と回答している割合がCBT形式では9%に対してPBT形式では26%と高く、逆に、この受験者の平均得点率は、CBT形式で27%であるのに対し、PBT形式では21%と低い値を示しており、これが相関に大きく影響していることが考えられる。

なお、高等学校での化学実験の頻度について、8択(1.しなかった 2.年に1回以上 3.学期に1回以上 4.月に1回以上 5.2週に1回以上 6.週に1回以上 7.毎回 8.化学を履修しなかった)で尋ねているが、1 ~ 3 (月に1回に満たない)と回答した割合の合計がA冊子69%, B冊子78%, 4と回答した割合がA冊子23%, B冊子16%であった。実験頻度が比較的低い傾向にあることもあり、高等学校での化学実験の頻度と得点率との相関係数は、形式別に見て、いずれもほぼ0で、無相関という結果であった。この結果から、高等学校での化学実験経験との関連を見ていくには、高等学校での化学実験経験が豊富な集団に対して調査を行う、具体的に経験した実験内容とそれに関連する問題の正誤との関係を見ていく、といったことをしていく必要があると考えられる。

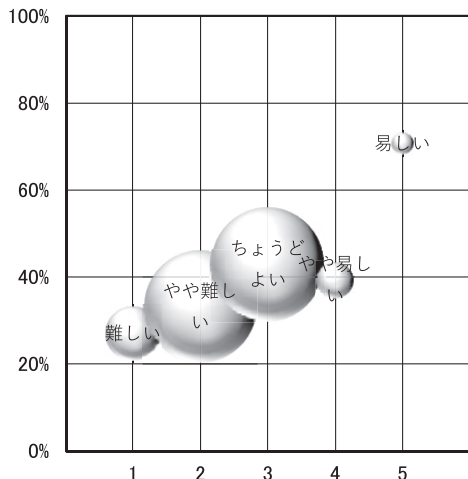


図7 質問と得点の相関 (A 冊子第2問:CBT)

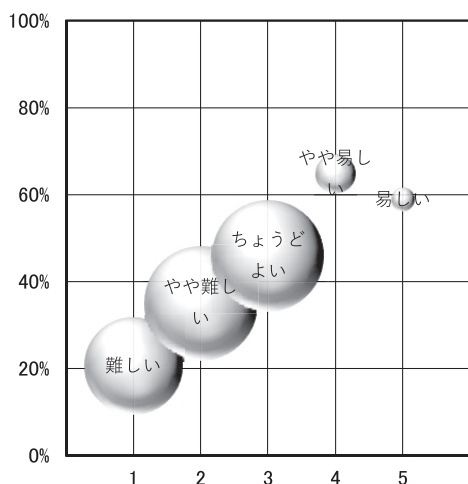


図8 質問と得点の相関 (B 冊子第5問:PBT)

## 附 記

本研究は、日本学術振興会科学研究費補助金基盤研究(A)「高大接続に資する多面的・総合的な学力評価・測定を行うための新たな技術的基盤の構築」(平成25～29年度, 課題番号25242016)及び基盤研究(A)「高大接続に資する多面的評価のためのタブレット端末を用いたCBTの総合的研究」(平成29～33年度, 課題番号17H00822)の成果の一部である。

## 後 註

iPad mini2, iBooksは、米国および他の国々で登録されたApple Inc.の商標である。

## 参考文献

- 安野史子(2017). 「高大接続を視野に入れたタブレットを用いる評価問題の試作—映像や動的オブジェクトを含む問題—」『大学入試研究ジャーナル』27, pp.71-79.
- 安野史子(2018). 「高大接続を視野に入れたタブレットを用いる評価問題の試作(2)—映像や動的オブジェクトを含む問題—」『大学入試研究ジャーナル』28, pp.155-162.

## 6 まとめと今後の課題

本研究では、映像や動的オブジェクトを含む化学のCBT問題について、PBTと評価がどのように異なるのかを探るため、PBT版に改作し、比較調査を行ったが、大問レベルで見ると、大部分の問題は、形式による難易度の違いが確認できなかった。小問レベルで見ると、現象や変化に関する明示の違いによって、差がある間が確認できた。そもそも、CBT形式に特化した問題を作題してきたこともあり、PBT形式へ書き換えがしにくい問題が多く、CBT形式とPBT形式では、測定している能力が変わってしまったものもあり、さらに個別に結果を精査する必要がある。映像やシミュレーションといった動的オブジェクトの働きと、評価したい能力との関係を明確にしていくことで、CBT問題の有益性は高まると考える。さらに、問題開発のみならずタブレットへ直接入力する解答方法の開発等、CBTに関する総合的研究として進めていく予定である。