

無回答および得点に基いた記述式問題の項目分析

—群馬県児童生徒学力診断テストにおける教科横断的検討—

石井 秀 宗*

1. 問題と目的
2. 方法
3. 結果
4. 考察
5. まとめ

1. 問題と目的

OECD（経済協力開発機構）が2000年から3年ごとに実施しているPISA（Programme for International Student Assessment）における2003年の結果において、日本の生徒の学力、特に読解力が著しく低下したということが報告された（国立教育政策研究所、2002、2004）、これは学力低下問題の議論に大きな影響を与え、文部科学省が2007年4月から、43年ぶりという全国一斉学力テスト（全国学力・学習状況調査）を実施するまでに至った。

PISAで明らかにされたことは、読解力の低下だけでなく、記述式問題に対する日本の生徒の無回答率の高さも指摘された（有元、2005など）。記述式問題における無回答率の高さは、平成15年度小・中学校教育課程実施状況調査結果（国立教育政策研究所教育課程研究センター、2005）など、他の調査やテストにおいても多数報告されている。平成19、20年度に実施された全国学力・学習状況調査の分析結果報告（国立教育政策研究所、2007、2008）においても、設問別の無解答率が報告されているように、無回答率が高いことを懸念する声は高い。

しかし、いずれの報告においても、設問ごとに無回答率の値を報告するか、記述式問題では無回答率が高いということを指摘している程度で、無回答をもたらす要因についての詳細な検討や、教科をまたがった横断的な議論はなされておらず、どのような記述式問題で無回答率が高くなっているのか、無回答と関連する要因にはどのようなものがあるかなどについての分析は、これまでほとんどなされていないのが現状である。

そのような分析を行った数少ない例として、石井（2007a）は、群馬県児童生徒学力診断テスト

* 名古屋大学大学院教育発達科学研究科准教授

のデータを分析して、記述式問題の無回答率の高さと関連する要因の検討を行い、物語の読み取りが楽しいと思う生徒とそうでない生徒を比べると、後者の方が記述式問題に対する無回答が多かったことを明らかにし、さらに、その傾向は小学校よりも中学校で大きいことなどを報告している(石井、2007b)。

これらの研究は、記述式問題に対する無回答に関連する要因を詳細に検討した研究として評価できるが、対象教科が国語に限定され他の教科については検討していないという課題があった。そこで本研究では、同じ群馬県児童生徒学力診断テストのデータを用いて、記述式問題に対する無回答に関連する要因を教科横断的に検討することを試みる。具体的には、群馬県児童生徒学力診断テストで収集された中学校3年生のテストデータを用いて、無回答と得点の観点から記述式問題を分類し、設問の内容や形式などについて検討する。無回答及び得点以外にも記述式問題を分類する観点はあり得るが、無回答と得点は、反応率や識別力を算出することで、テスト理論における項目分析の手法を用いて項目の特性を捉えることが可能であること、また、同テストが原則として1人の生徒に1教科のテストを割り当てる形式で実施され、教科を横断して共通に扱い得る適当な指標が他に見あたらなかったことなどから、この2つの観点に基づいて、記述式問題の分類を行うこととした。

なお、本稿では、「設問」「問題」「項目」という用語を文脈によって使い分けているが、いずれも「テストに含まれる最小採点単位の個々の設問」のことを意味するものとする。

2. 方法

(1) テストの概要

群馬県児童生徒学力診断テストは、群馬県教育委員会が2006年4月に群馬県全県の小中学校の児童生徒を対象に実施したテストであり、その目的は、児童生徒の学力の実態から各教科の今後取り組むべき課題をより明確にし、『[生きる力] 育成施策』にかかる『「確かな学力」ぐんぐんプラン』(群馬県教育委員会、2004)を推進していくための諸施策を提言するとともに、具体的な事業を展開することにあるとされている。なお、テストの作成は、群馬県総合教育センターの指導主事等により行われている。

同テストが実施された学年と教科は、小学校6年生の国語、社会、算数、理科、及び、中学校3年生の国語、社会、数学、理科、英語であり、基本的に各学校においては、いずれか1教科のテストが実施された。調査に参加した学校数は、小学校340校、中学校174校であり、郡部と都市部、北部と南部など地域の偏りがないように、教科の割り振りが計画された。テスト実施に要する時間は、小学校は45分、中学校は50分であった。

(2) 分析対象

群馬県児童生徒学力診断テストのうち、中学校3年生4教科の記述式問題41問(国語9問、社会16問、数学9問、理科7問)を分析対象とした。英語テストにも記述式問題はあったが、他の4教科が母国語で記述するのに対し、英語では外国語を書くという質的な違いが含まれることから、本研究では分離して扱うこととした。なお、各テストの全設問数は、国語17問、社会25問、数学30問、

理科32問であった。

分析対象のテストの受験者数は、国語3,892名、社会3,973名、数学3,860名、理科3,984名であった。また、当該教科の記述式問題すべてに回答した受験者数は、国語2,271名、社会1,191名、数学1,272名、理科1,956名であった。

(3) 分析方法

本研究では、記述式問題を無回答と得点の2つの観点から捉えることとしている。そこでまず、各項目について次の4つの指標の値を計算した。

無回答率：当該教科受験者のうち、その項目に無回答であった受験者の割合。

無回答識別指標：当該項目への回答の有無（有=1、無=0）と、当該項目を除いた同一教科内の他の記述式問題への無回答数の合計値との相関係数。

得点率：当該項目に回答している受験者における、項目得点（0～1）の平均値。

得点識別指標：当該教科の記述式問題すべてに回答がある受験者における、その項目の得点と、当該項目を除いた同一教科内の他の記述式問題に対する得点の合計値との相関係数。

無回答率は、値が高い（1に近い）ほど、その項目に無回答であった受験者の割合が高く、反対に、値が低い（0に近い）ほど、その項目に無回答であった受験者の割合が低いことを表す。

無回答識別指標は、値が高い（1に近い）ほど、その項目に無回答であった受験者が同一教科内の他の記述式問題にも無回答である傾向が強いことを表し、反対に、値が低い（0に近い）ほど、その項目に無回答であった受験者が同一教科内の他の記述式問題に無回答である傾向が弱いことを表す。

得点率は、値が高い（1に近い）ほど、当該項目に回答している受験者において、その項目に正答する受験者の割合が高く、反対に、値が低い（0に近い）ほど、その項目に正答する受験者の割合が低いことを表す。なお、いくつかの項目では、部分正答として部分点を与えている。

得点識別指標は、値が高い（1に近い）ほど、当該教科の記述式問題すべてに回答がある受験者において、その項目に正答した受験者が同一教科内の他の記述式問題にも正答する傾向が強いことを表し、反対に、値が低い（0に近い）ほど、その項目に正答した受験者が同一教科内の他の記述式問題に正答する傾向が弱いことを表す。

なお、得点率の算出にあたり、当該項目に回答している受験者に限定した理由は、当該項目に無回答であれば当然得点は0となり、得点率の中に無回答率の影響が入ってしまうのを避けるようにしたためである。同様に、得点識別指標の算出にあたり、当該教科の記述式問題すべてに回答がある受験者に限定した理由は、記述式問題に対する得点の合計値の中に無回答の影響が入ってしまうのを避けるようにしたためである。

次に、これら4つの指標を用いて、4教科の記述式問題41問を対象とし、階層的クラスター分析（ward法、平方ユークリッド距離、標準化得点）を行って、設問の分類を行った。そして、各クラスターに分類された設問の構成や解答類型などを検討し、設問の特徴と、無回答及び得点との関連を考察した。なお、データ解析にはSPSS ver. 15.0を用いた。

3. 結果

階層的クラスター分析の結果から、41個の記述式問題を8クラスターに分けることとした。階層的な分析では隣接するクラスターほど指標値の類似性が高いことから、クラスターの順序は保存して分析を進めることにした。

各クラスター及び全体の項目数、4つの指標の平均値と標準偏差を表1に示す。表1における標準偏差には、項目数が少ないことを考慮して、不偏分散の平方根ではなく、標本標準偏差を用いている。全41項目の4指標の平均値（標準偏差）は、無回答率0.195（0.124）、無回答識別指標0.505（0.111）、得点率0.539（0.200）、得点識別指標0.307（0.085）となっている。

また、各クラスターにおける4指標の傾向を概観するために、各指標の値を5段階に分割して表したものが表2である。表2では、全体平均+1.5SD以上の値を「特に高」、全体平均+0.5SD～全体平均+1.5SDの範囲の値を「高」、全体平均-0.5SD～全体平均+0.5SDの範囲の値を「中」、全体平均-1.5SD～全体平均-0.5SDの範囲の値を「低」、全体平均-1.5SD未満の値を「特に低」と表現している。それぞれのクラスターにおける各指標の平均値の傾向は次のようにまとめられる。

表1 各指標の標本平均値と標本標準偏差

cluster	無回答率	無回答 識別指標	得点率	得点 識別指標
1 4項目	0.225 .017	0.577 .023	0.263 .075	0.339 .032
2 2項目	0.505 .013	0.553 .003	0.161 .041	0.257 .053
3 8項目	0.313 .053	0.600 .017	0.451 .140	0.223 .065
4 3項目	0.232 .077	0.504 .078	0.543 .125	0.458 .019
5 13項目	0.182 .053	0.516 .057	0.643 .120	0.323 .036
6 7項目	0.059 .021	0.467 .058	0.749 .120	0.248 .062
7 2項目	0.025 .003	0.147 .004	0.483 .008	0.459 .023
8 2項目	0.041 .013	0.349 .046	0.458 .004	0.341 .003
全体 41項目	0.195 .124	0.505 .111	0.539 .200	0.307 .085

上段：平均値、下段：標準偏差。

表2 クラスターごとの各指標の傾向

cluster	無回答率	無回答 識別指標	得点率	得点 識別指標
1	中	高	低	中
2	特に高	中	特に低	低
3	高	高	中	低
4	中	中	中	特に高
5	中	中	高	中
6	低	中	高	低
7	低	特に低	中	特に高
8	低	低	中	中

特に高：平均+1.5SD～

高：平均+0.5SD～平均+1.5SD

中：平均-0.5SD～平均+0.5SD

低：平均-1.5SD～平均-0.5SD

特に低：～平均-1.5SD

クラスター1は、無回答率は0.2強と中程度だが、無回答識別指標は0.6弱と高い。得点率は0.3弱と低く、得点識別指標は0.3強と中程度である。

クラスター2は、無回答率は0.5程度と特に高く、無回答識別指標は0.6弱で中程度。得点率は0.2弱と特に低く、得点識別指標も0.3弱と低い。

クラスター3は、無回答率は0.3程度と高く、無回答識別指標も0.6程度と高い。得点率は0.5弱と中程度であるが、得点識別指標は0.2強と低い。

クラスター4は、無回答率は0.2強と中程度であり、無回答識別指標も0.5程度と中程度。得点率も0.5強と中程度だが、得点識別指標は0.5弱と特に高い。

クラスター5は、無回答率は0.2弱と中程度であり、無回答識別指標も0.5程度と中程度。得点率は0.6強と高いが、得点識別指標は0.3強と中程度である。

クラスター6は、無回答率は0.1未満と低いが、無回答識別指標は0.5弱で中程度。得点率は0.7強と高いが、得点識別指標は0.2強と中程度である。

クラスター7は、無回答率は0.1未満と低く、無回答識別指標も0.1強と特に低い。得点率は0.5弱と中程度だが、得点識別指標は0.5弱と特に高い。

クラスター8は、無回答率は0.1未満と低く、無回答識別指標も0.3強と低い。得点率は0.5弱と中程度であり、得点識別指標も0.3強と中程度である。

各クラスターに分類された項目の4指標の値、教科と設問番号、設問の概要を表3に示す。表中の括弧内の字数の目安は、回答欄の大きさ及び解答類型表に記載された解答例などの内容から推定した値であり、実際のテストにおいては、とくに字数制限はされていない。なお、クラスター内の項目の配列は、表の見やすさを考慮して、教科順、設問番号順に並べ替えてある。

表3 クラスターの構成

cluster	無回答率	無回答識別指標	得点率	得点識別指標	設問	概要
1	.248	.600	.264	.313	社23	3つの資料から読み取れることを記述する。ただし、個数の指定無し（60字程度）
1	.208	.550	.233	.388	理13	実験結果に対する予想の理由を記述する（50字程度）
1	.209	.558	.174	.307	理14	燃焼の前後で質量が変化する理由を記述する(40字程度)
1	.234	.598	.381	.347	理22	与えられた用語を用いて現象を説明する（50字程度）
2	.517	.550	.120	.204	社20	3つのカードに書かれていることから共通点を考えまとめる（15字程度）
2	.492	.556	.202	.309	数30	与えられた事実（証明）から、さらに明らかになることを記述する（20字程度）
3	.320	.588	.469	.293	社11	3つの図から情報を読み取り、象徴的な標語を作成し記述する（30字程度）
3	.393	.576	.549	.103	社12	図が表していることを表現する見出し語を作成する（20字程度）
3	.326	.596	.632	.215	社13	（資料に基づき）省エネ対策として自分にできることを考えて記述する（30字程度）
3	.271	.582	.416	.221	社16	南蛮貿易を盛んに行った理由を、資料に基づいて記述する（20字程度）
3	.264	.621	.134	.158	社24	2つの資料から読み取れることを記述する(40字程度)
3	.223	.617	.409	.215	社25	4つの資料から読み取れることを記述する。ただし、個数の指定無し（60字程度）
3	.351	.596	.552	.271	数04	別の式での考え方を図で示す
3	.356	.620	.445	.309	数29	数学的証明を記述する（10行程度）
4	.338	.599	.617	.464	数11	2変数の関連を表す式の立式の理由を、式や言葉を使って記述する（40字程度）
4	.200	.407	.646	.478	数15	会話を読み取り、状況をグラフ化する
4	.157	.506	.367	.433	理06	実験結果に対する予想とは異なる意見にどう答えるかを記述する（50字程度）
5	.143	.479	.509	.305	国08	「人間の外的な身体能力の拡張」を、具体例を挙げて説明する（30字程度）
5	.206	.585	.717	.275	国13	文章中で筆者が断言していることを、当該段落の語句を使って記述する（30字程度）
5	.237	.565	.846	.283	国14	本文中の内容を当該段落の語句を使ってまとめる（30字程度）
5	.135	.553	.544	.377	国16	要約の空欄に当てはまる内容を書く（30字程度）
5	.147	.428	.701	.293	国17	この文章を読んで共感したことを1つ取り上げ、感想や意見を書く（150字程度）
5	.241	.490	.658	.292	社05	自分でもできる防災対策を2つ挙げる(20字程度×2)
5	.081	.498	.384	.299	社08	図から読み取れることを記述する（20字程度）

5	.136	.505	.616	.337	社19	絵から読み取れる事柄を記述する (60字程度)
5	.230	.580	.753	.390	数02	式での考え方を図で示す
5	.134	.444	.718	.310	数09	答えの予想についての理由を、式や言葉を使って記述する (30字程度)
5	.201	.553	.521	.349	数13	2変数の関連を表すグラフを選んだ理由を記述する (40字程度)
5	.253	.435	.653	.358	理18	条件下の状況をグラフに書き加える
5	.221	.599	.739	.335	理25	条件を変えたときの様子の変化を記述する(20字程度)
6	.019	.377	.944	.251	国07	まとめ図の空欄に当てはまることばを本文から抜き出して記述する (10字程度)
6	.090	.525	.767	.301	国09	「人間がもつ内的な身体能力の衰え」を、具体例を挙げて説明する (30字程度)
6	.062	.534	.805	.256	国15	要約の空欄に当てはまる内容を書く (30字程度)
6	.067	.469	.554	.252	社06	図から読み取れることを記述する (20字程度)
6	.045	.478	.611	.210	社07	図から読み取れることを記述する (20字程度)
6	.051	.385	.792	.128	数07	表から読み取れる2変数間のきまり(関連)を記述する (20字程度)
6	.077	.498	.770	.340	理08	実験結果に対する予想の理由を記述する (40字程度)
7	.028	.143	.475	.436	社10	2つの図から情報を読み取り、条件に合う国を選択しその理由を記述する (20字程度)
7	.022	.151	.490	.481	社18	日米修好通商条約のどの条文が、日本にとってどう不平等な内容なのか記述する (20字程度×2個)
8	.028	.395	.462	.337	国01	筆者が読者に伝えようとしている内容を3つまで記述する (30字程度×3個)
8	.054	.303	.454	.344	社02	与えられた用語を用いて、選択理由を記述する (30字程度)

4. 考察

各クラスターにおいて、設問の構成や解答類型などにより把握される設問の特徴と、無回答及び得点との関連について考察し、最後に総合的な考察を行う。

(1) クラスター1

クラスター1は、社会1項目、理科3科目からなり、無回答率が中程度であるにも関わらず、得点率が低いのが特徴である。クラスター4や5も無回答率は中程度であるが、得点率は中か高であり、これらのクラスターと比較して、クラスター1は回答しても得点を得るのが難しい項目の群となっている。理科の項目は、実験での変化や結果についての理由を記述する設問などで、このような項目は、回答を書くことはある程度できるものの、難易度が高くなる傾向にある項目であると考えられる。

(2) クラスター 2

クラスター 2 は、社会 1 項目、数学 1 項目からなり、無回答率が 5 割程度と特に高く、得点率が特に低いのが特徴である。

社会の項目は、ある学習課題に従って作成されたという、3 つのカードに書かれている「主なできごと」を時代の流れとして見ていき（筆者注：波線部は波線も含めて実際のテストで使われた言葉である）、その学習課題を「どのようにして、…したのでろうか」のようにまとめるという課題である。カードの内容は次の通りである。なお、前段階の項目で、カードを時代順に並べ替えるという設問がある。

カード A：応仁の乱はじまる。日本に鉄砲が伝来する。南蛮貿易がさかんになる。織田信長が足利義昭を京都から追放する。

カード B：ペリーが浦賀に来航する。日米修好通商条約が結ばれる。薩長同盟が成立する。徳川慶喜が大政奉還を宣言する。

カード C：文永の役がおこる。弘安の役がおこる。永仁の徳政令がだされる。鎌倉幕府が滅びる。

この設問の解答例は、「どのようにして、幕府は滅亡（交代）したのでろうか」となっている。

数学の項目は、前段階で平行四辺形とその対角線からなる辺や角の性質をまとめ、対角線の交点を通る直線を引くことによって長さの等しい線分の組がいくつかできることを証明し、それらを受けて当該項目では、さらにどんなことが明らかになるかを答える問題である。なお、この項目は、数学テストの最後に配置されている。解答例は、「平行線は対角線の交点に対して点対称である」または「平行四辺形の対角線の交点を通る直線は平行四辺形の面積を二等分する」である。

いずれの設問も約半数の生徒は回答しなかったことから、設問の配置を考慮したとしても、設問の設定と解答例との乖離が大きかったことが推察される。解答を知っていれば設問文から解答を類推することができるかもしれないが、解答を知らない者にとってはその推測は困難なものとなり、回答を書いた生徒にあっても、得点率は社会の項目で 1 割、数学の項目で 2 割と低くなっている。生徒にとっては、設問の意味がよく理解できず、何を書いたらよいかよく分からない項目であったと言えよう。

(3) クラスター 3

クラスター 3 は、社会 6 項目、数学 2 項目からなり、無回答率が高く、また、得点率が中程度であるにもかかわらず得点識別指標が 0.223 と最も低いのが特徴である。クラスター 3 に含まれる項目に無回答である生徒は全体の 3 割程度もあり、回答を書くことがまず相対的に難しい項目である。また、得点識別指標が低いことから、その項目で得点できること（できないこと）と、同一教科の他の記述式問題で得点できること（できないこと）との関連が弱く、他の項目とは異質な能力を測定している可能性が考えられる。設問内容は、複数の資料を見て情報の読み取りを行ったり、読み取れることを見出しや標語で表現する、式で表されていることを図で説明するなどであり、統合力を必要とする項目の群であると考えられる。

(4) クラスター 4

クラスター 4 は、数学 2 項目、理科 1 項目からなり、無回答率は中程度であり、また、得点率が中程度である一方、得点識別指標は特に高いのが特徴である。当該のテストで測定されている能力が高い生徒は得点できるが、その能力が低い生徒は回答できないか、回答したとしても正答には至りにくく、その教科の記述式問題が全体としてどのような能力を測定しているかを象徴するような項目の群であると考えられる。数学の項目は、水槽に水を入れる問題で、時間と水位の関連を表す式の立て方を、式や言葉を使って説明する項目と、ウサギとカメの丘登り競争の様子を表した会話文から競争の様子を読み取り、それを横軸に時間、縦軸に距離を取ったグラフに表現する項目である。理科の項目は、燃焼前後の質量の比較に関する問題で、前段階で実験結果の予想を立て、当該設問では、その予想とは異なる意見と理由を言う人に対して、自分の意見を説明するというものである。いずれの項目も、当該教科の内容をよく理解していること、及び、それを正しく表現・説明できることが必要であり、いわゆる深い理解と表現力が求められる設問であると考えられる。

(5) クラスター 5

クラスター 5 は、得点率が高い以外はどの指標も中程度であり、無回答や得点の指標自体に顕著な傾向があるわけではないが、4 教科いずれの項目も複数個ずつ含まれていることが特徴である。相対的に国語の項目が多く含まれており、文中で書かれていること具体例を挙げる、文章を要約する、図などから読み取れることを書く、考えを図に示す、前段階での答えの理由を書くなど、設問形式は様々であるが、統合力(クラスター 3)や、深い理解と表現力(クラスター 4)を必要とする項目に比べ、単要素的な記述力、すなわち、文章の要約、具体例の記述、理由の記述、図表の読み取り、図示などの能力を個々に測定している項目群であると考えられる。全教科にまたがるクラスターが構成されたことから、これらの要素は教科を越えて共有される共通要素である可能性が示唆される。

(6) クラスター 6

クラスター 6 も、クラスター 5 と同様、4 教科いずれの項目も複数個ずつ含まれているが、無回答率が低いのが特徴である。設問形式は、言葉の抜き出しや当てはめ、図表からの情報の読み取りなどであり、クラスター 5 に含まれる項目よりも記述量が少なく回答負荷が小さい傾向にある。それゆえ、無回答率が低くなっているのだと考えられる。

(7) クラスター 7

クラスター 7 は、社会 2 項目からなり、得点率が中程度である一方得点識別指標が特に高いというクラスター 4 で見られた傾向を持つが、クラスター 7 は無回答率が低く、無回答識別指標も特に低くなっているのが特徴である。条件にあうものを選択するという前問と入れ子構造になっていて、前問に正答した者のみを対象として解答類型が設定されていたため、前問で誤答した 4 割弱の生徒は分析から除外されている。それゆえ、前問に正答できるような 6 割強の生徒だけが分析対象となった当該項目においては、無回答率は低く、無回答識別指標も特に低くなっているのだと考えら

れる。設問の内容は、2つの図から情報を読み取り、条件に合う国名を選択してその理由を記述する項目と、日米修好通商条約の条文の中から不平等条約であるものを選択し、その理由を記述する項目である。得点識別力が高いことから、クラスター4と同様、いわゆる深い理解をしていることが求められる設問であると考えられる。しかし、4割近くの生徒が除外されていることから、解答類型の作成に改善の余地が認められる。1つの方法として、前問で誤答した場合でも、当該項目の回答を「無回答」と「その他の回答」の2つに分類することが考えられる。

(8) クラスター8

クラスター8は、国語1項目、社会1項目からなり、無回答率が低く、得点率も中程度であるが、得点識別指標も中程度に留まっているのが特徴である。国語の項目は、問題文を読んで、筆者が伝えようとしていることを3つまで記述するという設問であり、条件を満たすものを3つ書いて完全正答、それ以外の場合は部分正答などとなる。社会の項目は、気候のグラフを見てどの地域のものかを記号で選択し、その理由を記述するという設問であり、記号と理由の双方を正しく書いた場合のみ正答となる。これらの項目においては、少なくとも1つのこと、または選択記号を書くことにはさほど困難は無いことから、無回答率は低くなっていると考えられる。しかし、完全正答に至るには複数の条件が設定されていることから、得点率が中程度、得点識別指標も中程度に留まったものと考えられる。

(9) 総合考察

上述した各クラスターに含まれる項目の主な特徴をまとめたものを表4に示す。総合考察として、クラスター分析に用いた指標値に加え、表4に示した項目の特徴も考慮して、記述式問題の諸特徴を整理してみることにする。

表4 各クラスターに含まれる項目の主な特徴

cluster	特徴
1	理科の問題が多く、実験での変化や結果について理由を記述する項目
2	設問の意味が不明確であり、解答類型と生徒の回答との対応が悪い項目
3	社会の問題が多く、統合力などが必要な項目
4	正答に至るには深い理解と表現力を必要とする項目
5	全教科にまたがり、諸々の記述力を単要素的に測定している項目
6	全教科にまたがり、言葉の当てはめや情報の取り出しを測定している項目
7	前問と入れ子構造になっているという問題はあがるが、正答に至るには深い理解を必要とする項目
8	設問の形式上無回答とはなりにくい、完全正答に至るのは難しい項目

クラスター1とクラスター3は、それぞれ、理科、社会の項目が多いという特徴が見られる。これらのクラスターに含まれる項目は、当該教科のテストに固有な記述式問題の性質を持っていると考えられる。理科では、実験での変化や結果についての理由を説明する設問、社会では、複数の資

料を見て情報を読み取って統合し、標語などで簡潔に表現する設問がこれに相当している。

クラスター4とクラスター7は、当該教科の内容をよく理解していること、及び、それを正しく表現・説明できることが必要であるという特徴が見られる。いわゆる深い理解をしていることが求められる設問からなっており、総合的な能力を測定することに向いていると言われる記述式問題の例になると考えられる項目の群であると考えられる。

クラスター5とクラスター6は、文章の要約、具体例の記述、理由の記述、図表の読み取り、図示、言葉の当てはめ、情報の取り出しなど、記述力に含まれる諸々の能力を単要素的に測定する項目の群であると考えられる。いずれのクラスターにも全教科の項目が含まれたことから、これらの要素は教科を越えて共有される共通要素であり、それに基づいて設問が作成されたと推察される。

クラスター2、クラスター8、クラスター7は、設問形式や解答類型の特徴から、他の設問とは異なるクラスターを形成した項目の群であると考えられる。クラスター2の項目は、生徒にとっては設問の意味がよく理解できず、何を書いたらよいのかよく分からない項目であったと考えられ、設問文、解答類型の双方について検討が必要であると思われる。とはいえ、回答負荷が大きい項目にもとにかく何か書いてみようという姿勢を持たない生徒が半数近くいるということを明らかにした意義は大きいと考えられる。

無回答率と設問形式との対応を見てみると、回答負荷が小さい項目では無回答率は1割未満と低く（クラスター6、7、8）、要約や具体例を挙げること、理由を書くことなどの設問になると、1～2割程度の無回答率となり（クラスター1、4、5）、複数の資料からの情報の読み取りや、その情報を簡潔に表現する設問では、無回答率は3割程度と高くなる（クラスター3）。また、設問の意味がよく理解できず、何を書いたらよいのかよく分からない設問になると、5割程度の生徒が無回答となる（クラスター2）、ということが理解される。

5. まとめ

本研究では、無回答と得点の観点から記述式問題の分類を行い、記述式問題の特徴を教科横断的に検討することを試みた。分析の結果、8つのクラスターが形成され、記述式問題で測定しようとする能力について、文章の要約、具体例の記述、理由の記述、図表の読み取り、図示、言葉の当てはめ、情報の取り出しなど、教科を越えて共有されるいくつかの共通要素が存在し得ること、測定しようとする要素の質や量の違いによって無回答率に差が生じること、設問形式や解答類型など形式的な要因も無回答率に影響することなどが明らかにされた。

【謝辞】

本研究は、群馬県総合教育センターとの共同研究を継続した科学研究費補助金研究（若手研究（B））の一環としてなされたものです。当時、同センターの所長をされ、現在、県立高崎女子高等学校校長の飯野真幸先生、同主任指導主事で、現在、高崎市立箕郷東小学校校長の岡野健先生に深謝致します。問題開発に携われた先生方、テストの実施及び評価にあたられた先生方にも感謝致します。また、テストを受験した生徒の皆さんに御礼申し上げます。

【文献】

- 有元秀文(2005)。PISA読解力調査で分かったこと。公開シンポジウム報告書「中等教育段階における多面的教育測定 ―PISAを超えて」。東京大学大学院教育学研究科教育測定・カリキュラム開発講座、16-29。
- 群馬県教育委員会(2004)。平成16年度教育の指針。
- 石井秀宗(2007 a)。記述式問題における無回答に関連する要因の検討 ―群馬県児童生徒学力診断テスト小学校6年生国語テストデータ分析の結果から。日本テスト学会誌、3、59-70。
- 石井秀宗(2007 b)。群馬県児童生徒学力診断テストデータの分析。東京大学大学院教育学研究科教育測定・カリキュラム開発(ベネッセコーポレーション)講座2006年度研究活動報告書。Sokutei Report、6、234-250。
- 国立教育政策研究所(編)(2002)。生きるための知識と技能 ― OECD生徒の学習到達度調査(PISA)2000年調査国際結果報告書。ぎょうせい。
- 国立教育政策研究所(編)(2004)。生きるための知識と技能2 ― OECD生徒の学習到達度調査(PISA)2003年調査国際結果報告書。ぎょうせい。
- 国立教育政策研究所教育課程研究センター(2005)。平成15年度小・中学校教育課程実施状況調査結果。
- 国立教育政策研究所(2007)。平成19年度全国学力・学習状況調査：設問別調査結果。
- 国立教育政策研究所(2008)。平成20年度全国学力・学習状況調査：設問別調査結果。

An Item Analysis of Open-Ended Questions Based on Nonresponses and Scores: A Cross-Subjects Study on the Achievement Test for Middle School Students in Gunma Prefecture

Hidetoki Ishii*

It is often worried that the proportion of nonresponse for open-ended question is high in Japanese middle school students. In this paper, using the data of the Achievement Test for Middle School Students in Gunma Prefecture, forty one open-ended questions in four subjects, that is, language, social science, mathematics and natural science, were clustered by indices of nonresponses and scores and several characteristics of open-ended questions were considered.

As the results, eight clusters of open-ended questions were conducted. And it was found that summarization of continuous texts, writing examples, writing reasons, reading of non-continuous texts, illustration of perception, choosing words and retrieving information could be common abilities in open-ended questions across four subjects. It was also found that the proportion of nonresponse was varied by the quality and quantity of the abilities that the item measured. Moreover, it was observed that item forms and construction of answer rubrics affected the proportion of nonresponse for open-ended questions.

* Associate Professor, Graduate School of Education and Human Development, Nagoya University