

## QQ Test の開発と試行

—探究学習の成果を測定するための指標づくりを目指して—

池田 文人, 岩間 徳兼 (北海道大学), 成田 秀夫 (大正大学)

2022 年度から施行される高等学校学習指導要領では「探究」の名が付く科目が 7 つ新設される。大学における研究力を高めるためにも、高校での探究学習の成果を大学入試において適切に評価することは重要な課題である。探究学習を通じて知識・技能を論理的思考により活用することが目指されるが、研究にとって真に重要なことは、リサーチ・クエスチョンとして科学的にものを問う能力である。しかし問うこと（質問）を評価する手法は見当たらない。そこで本研究では論理的思考を問う力を測定する QQ Test を開発し、探究活動に取り組む約 200 名の高校 1, 2 年生に試行した。本テストが各高校の特徴を識別する能力を有し、一般の学力テストとは異なる能力を測定しうることが示唆された。しかし質問数が少ないことが課題であり、質問の仕方の学習支援から取り組む必要がある。

キーワード：探究学習、リサーチ・クエスチョン、質問力、テスト開発

### 1 はじめに

高等学校学習指導要領の改定に伴い、2022 年度から「探究」の名前がつく科目が 7 つ新設される。その中で「総合的な探究の時間」は必修科目に指定される。高校での「探究的な学び」の成果を大学入試において適切に評価することは、高大接続の観点からも重要な課題である。しかし現在の高校教育では、学びの基本となる知識・技能が教科教育の中で完結してしまい、探究学習の中で活用できていないという指摘もある（安彦, 2016）。この課題を解消するために、知識・技能の習得と探究とをつなぐ知識・技能の活用を挟む学習モデルが提案されている（同）。ここでの活用で求められている能力は、知識（規則や命題）を同種の事例に適応したり、ある事例を既習の規則や命題を活用して解決したりするという演繹推論の能力である。

しかし、実際の科学的な研究においては、多様で雑多な具体的な事象やデータから、統計学に基づいて厳密かつ客観性の高い規則を導くという帰納推論の能力が求められる。一方、観察されたデータ（結果）からその原因を推定するという仮説推論（ベイズ推論）は、生じている問題事象からその原因を推定し解決する能力として重要であり、また AI が行う重要な推論の 1 つとして AIED (Artificial Intelligence in Education) においても重要性が高まっている（Holmes, etc. 2019）。高校での探究学習が大学での研究や社会での問題解決、さらには AI 時代に求められる教育につながるためには、帰納推論・演繹推論・仮説推論の力がバランス良く育成される必要がある（池田, 2018）。

一方、探究ひいては研究の質を左右するのはリサーチ・クエスチョン、すなわち「科学的問い」である。しかし、現在の教育において科学的な問い方を学ぶことはなく、したがってそれが評価されることもない（池田, 2018）。従来の試験は問題が与えられ、学習者は学習した知識が正しいことを前提に演繹的にその解決策を考え、正解と比較されることによって評価される。ここで前提としている知識の多くは有限回の実験や観察を通じて帰納推論により得られたものであり、正しいことが保証されているわけではない。したがって、知識が形成される帰納推論を疑問問う能力が探究ひいては研究において求められる。と同時に、演繹推論の論理性を批判的に問うことも、厳密な科学的な研究では求められる。さらに、得られたデータからそれが生じた原因を特定するという仮説推論も探究学習や科学的研究においては重要である。しかし原因は多様であり、持っている知識量によっても推定できる原因の数は変わってくる。したがって推定した原因がどのくらいの確率で確からしいのかを疑い、問う能力が求められる。つまり、帰納推論・演繹推論・仮説推論の力には、それぞれの推論そのものを批判的に捉えて問う能力が求められる（Ikeda, 2021）。

そこで、本研究では 3 つの論理的推論について問う力を Question-Intelligence (QI : 質問知能) と呼び、これを測定する指標を Question-intelligence Quotient (QQ : 質問知能数) と名づけ、QQ を測定する課題と採点基準を開発し、これを QQ Test と名づけ、試行した結果に基づきその妥当性と信頼性を検証した。現代において能力というのはとても大きな概

念として扱われている (松尾, 2015)。知識や技能に代表される認知能力だけでなく、態度や心理的側面を含んだ非認知能力までを範疇とすることが多く、能力・資質、コンピテンシー、リテラシーなど、能力やその側面を表現する用語は多岐に渡る。それは、学習を通して培われる力であるところの学力についても同様である。よく知られるように、文部科学省はそれを構成する三つの要素として「知識・技能」、「思考力・判断力・表現力」、「主体的に協働的な学びへ向かう態度」を示している。

ところで、学力を育む教育活動においては重要な要素の一つとして評価、すなわち、言葉や数値により学習者の熟達や到達の程度を表すことが行われる (中島, 2018)。特に、入学者選抜における評価は、測定、つまり、個人特性としての学力に対して数値を対応づけるという意味合いが強い。そして、学校等での教育場面や入学者選抜において多岐に渡る学力を評価もしくは測定するために、客観式テスト、論述式テスト、面接試験などが利用されている。

上記のような能力、学力とその評価、測定という視点から捉えると、QQ Test は単純な知識・技能に止まらず高次の能力、学力の三要素でいえば、「思考力・判断力・表現力」を対象とし、論述式テストによってその測定を試みるものである。そして、試験問題として3タイプの課題を設計しテストを行い、論理性の点から思考力・判断力を、問う相手と問うための表現形式の点から表現力を評価するように採点基準を設定して採点した。

本研究は、高校における探究学習において3つの推論力と科学的に問う力の育成を促すために、QQ Test を大学入試に導入することを目指す。個々の学校教育を変えることは多大な労力と時間を要するが、大学入試が変わればその変化に対応する教育も変わっていくと考えているためである。それだけ大学入試が高校教育へ与える影響は大きいと考えているため、大学での研究や社会での問題解決に有効な大学入試に変えていく必要がある。

## 2 質問の重要性

文部科学省による 2020 年度学校基本調査によれば、2020 年度の短大を除く大学進学率は過去最高の 54.5% となった。トロウ・モデルによれば、大学進学率が 50% を超えた国の大学教育の役割は、エリート・指導者育成から、産業社会に適応しうる全国民の育成への転換を迫られている (Trow, 1961)。しかしながら大学に期待される研究の役割は変わることはな

く、文部科学省が 2020 年度から大学の教育・研究にも導入した客観的な共通指標に示されるように、国際的に質の高い研究が求められており、そのための教育の質の向上も問われている。こうした質の高い研究活動にとって重要な役割を担うのが、リサーチ・クエスチョンとしての「質問」である。

相対性理論を生み出したアインシュタインはその著書の中で、質問はその解決より本質的だと書いている (Einstein and Infeld, 1967)。また、人文社会科学に構造主義をもたらしたレヴィ=ストロースは、科学者とは正しい質問ができる人だと言っている (Young, 2014)。自然科学や人文社会科学において、リサーチ・クエスチョンは研究の本質であろう。

一方、高校教育では 2019 年度入学者から新科目「探究 (総合的な探究の時間)」が先行導入された。その平成 30 年に告示された学習指導要領に示された目標の一つとして、「実社会や実生活と自己との関わりから問いを見だし、自分で課題を立て、情報を集め、整理・分析して、まとめ・表現することができるようにする。」と記載されている。つまり、探究は、問いすなわち質問から始まる。また 15 歳児を対象とした国際的学力調査として有名な PISA では、科学的な探究を評価しデザインする科学的コンピテンシーとして、質問への科学的な言及が必要だとしており (OECD, 2019)、質問を科学的探究の核としている。

さらに、実社会に出てからの、他者への質問の良し悪しが、仕事上の問題解決の質と効率を左右することも指摘されている (池田, 1999)。すなわち、質問によって自分が直面している問題のことばかり書いたのでは問題解決に有益な回答を得ることはできず、問題が生じた経緯や背景に関する情報を質問の中で詳細に記述した方が問題解決に有益な回答を迅速に得ることができる。

以上のことから、質問力は、高校と大学、そして社会までを接続する、探究や問題解決に有効な能力だと言える。また、質問力は、学力の3要素を網羅するものである (池田, 2018)。本質的な質問をするためには質問の対象に対する深い理解と多様な観点が必要であるため、幅広く深い知識が必要である。また他者への質問には、他者に質問の意図を端的に伝えるための表現力とそのため技能が必要である。そして問題を解決するためにはどのようなことを質問すべきかを多面的かつ批判的、論理的に思考する力が必要であり、何を質問することが問題解決の本質であるかを判断する力も必要である。さらに、質問は好奇心から生まれ

るものであるため主体性につながり、他者への質問は多様な人々との協働的に学ぶ態度を促す。したがって、質問力は学力の3要素を網羅できる重要な能力である。

### 3 質問を評価する難しさ

18世紀を代表する啓蒙主義者であるヴォルテール  
の名言として、人を判断するにはその人の回答ではなく質問をもってせよ、という有名な言葉があるが  
(Voltaire, 2020)、質問を評価する取り組みや研究は見当たらない。その理由は、質問を評価することが困難だからである。

評価とは、「価値を測って判断する」ことである。従来のテストの場合、解答の価値は正解との近さだと言える。正解というものが仮定されていれば、正解と解答との相違によって、それらがどれだけ近いかわかるといえる。具体的な客観的に測ることができ、その近さでその価値を判断できる。

しかし質問の価値を定義することは難しい。例えば、正しい質問という正解を仮定し、受験者の質問の価値はこの正しい質問との近さとする。受験者が出した質問と、この正しい質問との相違を調べれば、評価できそうである。では「正しい質問」とはどのようなものであろうか？質問に含まれる知識や論理が正しいことが正しい質問であり、質問の価値だとしてしまうと、科学的探究に不可欠な、質問のもつ創造性が失われてしまう。誰も気づかなかった新しい視点から質問することは、科学におけるコペルニクス的転回 (Kant, 2007)、あるいはパラダイム・シフト (Kuhn, 1962) を生み出すからである。しかし逆に突拍子もない珍奇な質問だからといって創造的な価値があるとは言えない。その質問が現実的に検証可能であること、すなわち既存の知識を用いて解決できる必要があるからである。

### 4 質問の評価方法

では質問の価値を何に求めれば、質問を評価できるであろうか。質問そのものの知識や論理を評価しようとすると、新規さと珍奇さの板挟みになってしまう。そこで、質問の対象と質問との関係性に価値を求める。すなわち、質問の対象は数式や図表を含む広い意味での記号としての言葉によって表現される何らかの知識であり、その知識を成立させている論理の不備を質問によってどれだけ問えるかを、質問の価値とする。つまり個々の質問を個別に評価するのではなく、質問対象に対する質問群として、どれだけ質問対象である知識を成立させている論理の不備を問えるかを評価する。

質問対象となる知識を成立させている論理の不備を問うことは、その知識に代わる、より論理的に堅固な知識を探究できる可能性を示唆するものであり、創造性につながる。また、単体の質問を評価するのではなく、質問対象に対する質問群を評価することにより、各質問が単に奇を衒うだけの質問かどうかを判断できる。質問対象となる知識に対してすべての質問が論理的につながっていなければ、質問群として価値をもたないからである。

1章で論じたように、論理には帰納推論と演繹推論、そして仮説推論の3種類がある。本研究ではこれらの推論について被験者に質問を挙げさせ、その質問を評価する。帰納推論の論理を問う質問とは、個々の事例が一般的な規則を導くためのサンプルとして質的・量的に妥当であるかどうかを問うような質問である。演繹推論の論理を問う質問とは、規則の組み合わせ方に矛盾や飛躍がないかといった論理を問うような質問である。そして仮説推論の論理を問う質問とは、個々の事例の妥当性や、その原因を推測するための規則の妥当性、あるいは規則の組み合わせ方の妥当性などを問うような質問である。

### 5 QQ Test 問題の開発

このような3種類の論理の不備を問うような質問を受験者に挙げさせるテスト問題を開発する。このようなテスト問題としては、3種類の論理ごとにそれぞれに特化した題材を出す方法と、3種類の論理について同じ題材を用いて一連の設問によって挙げさせる方法とが考えられる。前者は各論理に特化した題材を考案することにより、期待する質問を受験者から引き出しやすい。その反面、論理ごとに題材を作成する必要があるため、テスト開発のコストは高くなるとともに、受験者への負荷も大きくなる。後者は、同じ題材を用いて3種類の論理について問えるため開発コストも受験者への負荷も低くできる反面、異なる論理に対する質問を受験者から引き出すためには設問に工夫が必要となる。今回は、テストの開発のコストも受験者への負荷も抑えられる後者の方法で問題を開発した。

まず受験者が質問する対象となる資料を提示する。この資料では作者が導いた何らかの主張（一般的規則）が提示される。この主張について、以下の3つの設問が提示される。

設問1：資料に提示された主張が導かれるまでのプロセスについて疑問に思ったことを作者への質問としてできるだけたくさん挙げさせる。この設問では作者の帰納推論における論理の不備を問う質問力を測定す

る。

設問 2：資料に提示された主張が正しいと仮定して、その主張によって引き起こされる結果までのプロセスについて疑問に思ったことを、自分が探究するための自分への質問としてできるだけたくさん挙げさせる。この設問では作者の主張から演繹推論によって考えられる結果までの論理を批判的に問う質問力を測定する。

設問 3：設問 2 における結果を実現する、あるいは防止するために何をすべきかを明らかにするために、探究しなければならない事柄を質問の形式でできるだけたくさん挙げさせる。この設問では結果を実現、あるいは阻止するために、仮説推論によって効果的な方策を考え出すまでの論理を批判的に問うような質問力を測定する。

表 1 4校の受験形態

	A校	B校	C校	D校
媒体	冊子	Web	Web	冊子
時間	100分	自由	自由	自由
実施日	3月8日	3月初旬	3月初旬	3月11日
監督	有	無	無	無
調べ	可	可	可	可

このような一連の3つの設問を伴った大問を3題開発した。それぞれの題材は、健康問題、環境問題、食糧問題であり、これらはいずれもSDGsとして世界中

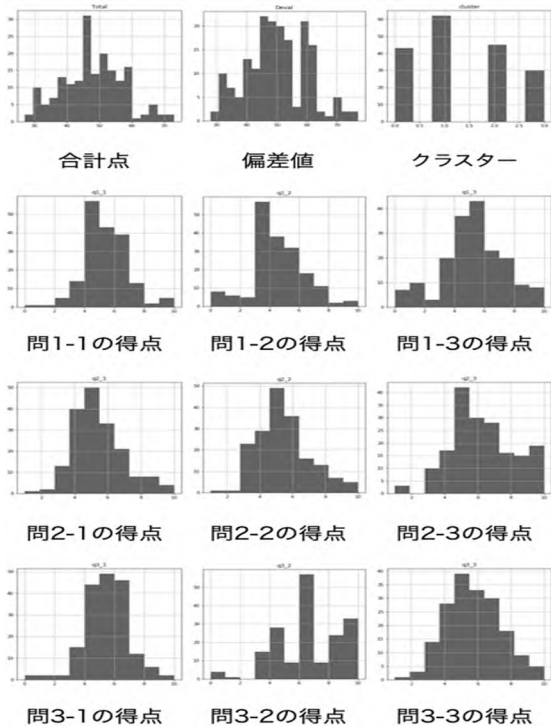


図 1 4校全体のヒストグラム

で取り組まれているものである。本テストの受験者は高校生であり、現在の高校における探究科目、すなわち「総合的な探究の時間」ではSDGsがテーマとして広く用いられていることから、広く高校生になじみがありながら、特定の教科・科目の知識に依存しない題材とした。健康問題に関する出題では、糖尿病に関する書籍の一部を読ませ、上述した3つの設問について質問を挙げる。環境問題に関する出題では、チッソの排出量に関するグラフや表と、そこから帰納推論によって引き出しうる主張とを読ませ、上述した3つの設問について質問を挙げる。食糧問題に関する出題では、日本の食糧自給率の変遷に関するグラフとそこから帰納推論によって引き出しうる主張とを読ませ、上述した3つの設問について質問を挙げさせる。

### 6 QQ Test の試行

近畿圏でSSH等において探究学習に精力的に取り組んでいる、私立高校1校と公立高校3校の主に理数科等に在籍する高校1,2年生であり、合計で196名の生徒が受験した。各高校の都合により、Web受験と冊子体での受験を選んでもらった。表1に4校の受験形態をまとめる。

Web受験では、大問3つを個々に分けて受験でき、約1週間の中で大問3つをすべて受験する。冊子体で受験した1校は高校教員の監督の下、100分で大問3つにすべて回答した。冊子体で行なったもう1校は受験場所と開始時間を指定し、全問終了し次第、解答用紙を提出してもらった。オンラインの場合も冊子体の場合も、受験者はインターネット等を自由に調べることができる。ただし、受験者同士で話し合うことはで

表 2 4校の基本統計量 (n=180)

設問番号	平均値	標準偏差	最小値	最大値
1-1	4.9	1.5	0	10
1-2	4.0	1.8	0	10
1-3	4.8	2.1	0	10
2-1	5.4	1.7	1	10
2-2	5.4	1.8	1	10
2-3	6.3	2.1	1	10
3-1	5.0	1.5	0	10
3-2	6.2	2.3	0	10
3-3	5.8	1.8	1	10
計	47.9	9.4	27	73

表3 各群の特徴

設問番号	高位群 n=43	中位群 n=62	低位群 n=45	特異群 n=30
1-1	6.1	5.1	3.9	4.3
1-2	5.1	4.6	2.6	3.3
1-3	6.5	5.4	3.0	3.9
2-1	6.4	4.8	4.4	6.8
2-2	6.3	4.9	4.3	7.0
2-3	8.4	5.6	4.4	7.6
3-1	5.9	5.0	4.3	4.8
3-2	7.4	6.4	4.6	6.4
3-3	7.7	5.6	4.4	5.4
計	60.0	47.5	35.9	49.6

きない。実際に使用した問題と解答用紙、および大問ごとの採点基準を <https://inquiry-evaluation.com/ギャラリー/qctest-2/> に掲載した。

### 7 試行の結果

全設問に回答した180名の受験者のデータを集計・分析の対象とした。4校の内訳は、A校が22名、B校が40名、C校が65名、D校が53名であった。

個々の質問は、前提・結論・根拠の論理性を批判的に捉え、何を問うかを適切に判断しているか（思考力・判断力）、それを質問の形式で適切に表現できているか（表現力）の2つの観点に基づき採点した。2章で述べたように質問は学力の3要素を網羅すると考えているが、3章で論じたように知識・技能を評価するためには正解を仮定する必要があり、質問のもつ創造性を損なってしまう。また主体性は問うという思考や行為の元になるものであるため、質問そのものとい

うよりも質問の数などにより評価されるものだと考える。協働的な学びの態度については、表現力を評価する観点として、他者へ質問することを意識した表現になっているかどうかを加えることにより、協働性を意識した。各設問につき思考力・判断力と表現力を各5点満点、合計10点満点とした。他者へ問うことは協働的問題解決力として重要であることから、表現力は論理性と同程度に重要だと考えた。また、論理性と表現力のそれぞれを優（5点）・良（3点）・可（1点）の3段階で評価し、点数にある程度のばらつきが出るようにした。採点者は2名で、2校ずつ担当した。両者は学習塾で20年以上にわたり国語の教師として授業を受け持ち、試験問題の作成や採点に携わってきた。採点に悩んだ場合は2名で協議して決定した。大問3問のそれぞれに設問が3問あるので、全体で合計90点満点である。

具体的な採点方法について説明する。例えば大問1の小問1の場合、思考力・判断力としては、筆者の主張とその前提およびそれらの関係について問えるかどうかを採点基準である。質問群としてこれらをすべて問えていれば5点、いずれか2つであれば3点、1つであれば1点である。この問題における筆者の主張は「個人を超えた社会関係を築くべき」というものであり、この主張に対して、「個人を超えた社会関係とは何か？」という質問であれば主張を問うていると判断した。また別な質問で「社会関係のベースが食と身体だけでよいのか？」という質問をしていれば前提を問うていると判断し、合わせて3点とした。表現力については同問の場合、自問自答ではなく他者を意識した質問形式となっていれば5点、自問自答的な質問であれば3点、質問の形式になっていなければ1点であり、一番得点の高い質問が1つでもあればそれを得点

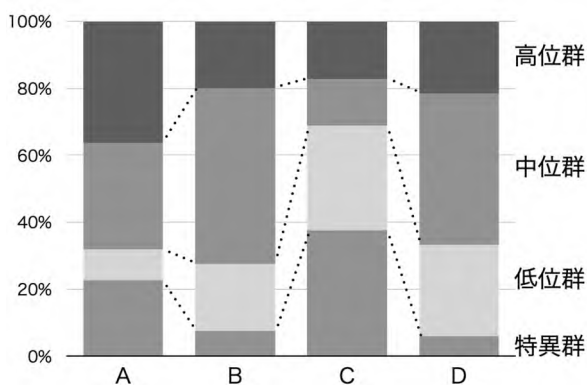


図2 各高校の各群の比率

表4 設問および合計点の間の相関係数

	1-1	1-2	1-3	2-1	2-2	2-3	3-1	3-2	3-3
1-1	1.0	0.3	0.4	0.2	0.2	0.3	0.2	0.2	0.3
1-2	0.3	1.0	0.5	0.1	0.2	0.3	0.2	0.1	0.2
1-3	0.4	0.5	1.0	0.1	0.1	0.4	0.1	0.2	0.3
2-1	0.2	0.1	0.1	1.0	0.4	0.5	0.2	0.1	0.2
2-2	0.2	0.2	0.1	0.4	1.0	0.4	0.2	0.1	0.2
2-3	0.3	0.3	0.4	0.5	0.4	1.0	0.2	0.1	0.3
3-1	0.2	0.2	0.1	0.2	0.2	0.2	1.0	0.2	0.2
3-2	0.2	0.1	0.2	0.1	0.1	0.1	0.2	1.0	0.2
3-3	0.3	0.2	0.3	0.2	0.2	0.3	0.2	0.2	1.0
計	0.6	0.6	0.6	0.5	0.5	0.7	0.5	0.5	0.6

とした。4 校を合わせた基本統計量を表 2 に、ヒストグラムを図 1 に示す。図中の「クラスター」については後述する。

まず、合計点を除く 9 設問の得点を用いて、K-means 法により受験者をクラスター分析したところ、4 クラスターの場合に意味のある分類が見られた。各クラスターの設問ごとの平均点に、合計点の平均を加えたものを表 3 に示す。

全ての設問について、得点が高い高位群、平均的な中位群、得点が低い低位群の他に、大問 1 の得点は低いが大問 2 の得点は上位群よりも高く、大問 3 は平均的な特異群に分かれた。高校ごとの、各群に該当する受験者数の比率を図 2 に示す。

次に主成分分析を行なった。第 1 主成分の寄与率は 32%程度と高くはないが、単調増加しており、比較的まとまりのあるテストだと言える。

最後に設問ごとの得点と合計点における相関分析の結果を表 4 に示す。設問間に関しては一番高くても 0.5 程度であった。この値は低い値ではないものの、ほとんどが 0.1 から 0.2 であり、ほとんどの設問間において相関は低い。しかし、合計得点と各設問との相関係数は総じて高い。

### 8 考察と今後の課題

4 校全体で見ると合計点や各設問の得点は、大問 3 設問 2 以外は、正規分布に近いものだった。一方、各高校の偏差値 (4 校全体) のヒストグラムを示した図 3 に示唆されるように、高校ごとに見ると得点分布に多数の断裂が見られ、得点ごとの人数も正規分布からは外れる。このことは各高校の特徴を反映

したものなのか、あるいは、本テストが高校を超えたより一般的な能力等を測定しているものなのか、被験者の層を拡大して調べる必要がある。

クラスター分析からは各高校の特徴が明らかになった。A 校は公立高校で自由な気風、そして C 校は中高一貫の私立高校であり、関心の偏った生徒がかなりの比率で存在し、特異群が B 校や D 校よりも多いことを裏付けている。B 校と D 校は優等生タイプの生徒が多い公立高校であり、中位群が半数近くを占めることを裏付ける。しかし、いわゆる学力試験の点数の高い生徒が、今回の QQ テストで高得点を取っているわけではない。学力試験の成績はそれほどでもないが、探究活動などでグループを引っ張っているような生徒が QQ テストで高得点であるケースが多かった。引き続き調査をして、QQ テストと探究活動との相関を検証する必要がある。

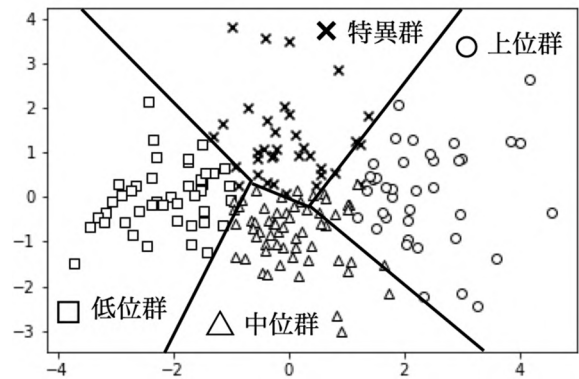


図 4 各群と主成分との関係

主成分分析では、第 1 主成分が約 0.32 であり、第 2 主成分を加えても 0.5 に満たない。そこで、クラスター分析と主成分分析の結果を紐付けて分析した。すなわち、第 1 主成分を x 軸、第 2 主成分を y 軸とした平面に受験者の得点を写像し、クラスターのラベルを付した様子を図 4 に示す。各群が比較的きれいに分別されているのがわかる。つまり、クラスター分析による 4 群が 2 軸 (2 次元) で表現できるということになり、測定される内容が 2 種類であることを意味する。1 次元目である第 1 主成分は高中低位群を分けている設問全体の得点の高低を反映したものだと考えられ、2 次元目である第 2 主成分は特異群を分けている指標だと考えられる。すなわち 2 次元目は探究活動や題材 (テーマ) への慣れもしくは前提知識だと推測される。このような、本テストで測定したい質問力以外の指標が入ることは望ましくない。慣れや前提知識に依存しない題材の選定を

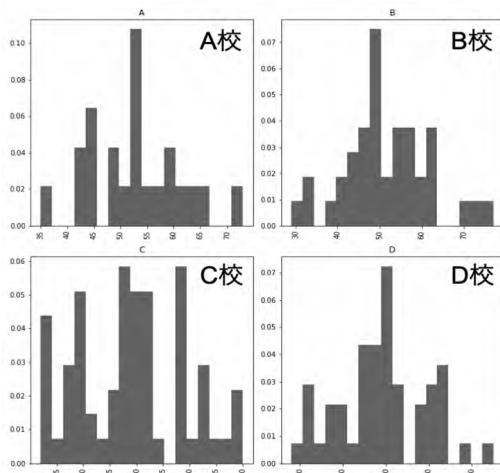


図 3 各高校の偏差値のヒストグラム

検討する必要がある。

相関分析において、設問間の相関は低く、設問と合計点との相関が高いことは、各設問が比較的独立して異なる能力を見ているように思われる。しかし、同じ小問は同じ論理を問う質問を挙げさせるため、相関は高くなることが想定されたがそうになっていない。設問ごとの質問数を調べたところ、各設問の平均質問数は 3 未満がほとんどであった。つまり質問 1 つあたりの得点はその設問の得点に与える影響が大きいということである。このことが設問間の相関の低さや学校ごとに見ると明確な分布が見られない原因の 1 つだと考えられる。設問あたりの 1 人当たりの質問数を増やさなければ、テスト問題の妥当性や信頼性を数量的に適切に検証することは難しい。

質問数が少ない原因としては 2 つ考えられる。1 つ目は、質問の数はその対象に対する関心の高さと比例すると考えられるので、質問数が少ないということは総じて関心が薄いということになる。しかし、今回の受験者は探究活動を重視している高校の、探究活動に取り組む機会の多い理数科等の生徒であることを考えると、受験者の関心が低いということは考えづらい。もう 1 つの可能性としては、質問の仕方が分かっていないということである。学年が上がるごとに子どもたちは質問しなくなり、その大きな要因は、質問することの恥ずかしさや分かっていないと思われることへの抵抗など、他者への意識だと言われる(藤井・山口, 2003)。しかし今回のようなテストでは、このような意識はあまり働かないであろう。そう考えると、そもそもどのように質問をしてよいか分かっていないことが原因だと考えられる。実際、受験した生徒から、どのように質問してよいか分からなかったという感想が多く聞かれたと、各高校の担当教員から報告を受けている。

今回のテストでは、ペーパーテストにおいてもオンラインテストにおいても、設問ごとに開始時間と終了時間を記入・入力してもらっている。質問の仕方が分からない場合、以下の二つの可能性が考えられる。所要時間が短く、質問も少なく、得点も低い場合。そして逆に質問の仕方を考えるのに時間を要したために、所要時間が長く、質問はそれなりに多く、得点もそれなりに高い場合である。適切なテストでは、適正な時間で適正な質問数と点数が取れることが望ましい。所要時間と質問の質と数の検証が必要である。

適正な時間で適正な質と量の質問をしてもらうためには、3 種類の論理を問う適切な質問の仕方を教

示する必要がある。すなわち、帰納推論の不備を問う質問にはどのようなものがあるか、演繹推論の場合はどうか、仮説推論の場合はどうか、といった具体例を示し、日頃からそのような質問をする機会を増やすような支援が不可欠である。このような質問の事例集を作成するとともに、適切な質問の仕方を誰もがいつでも学習できるような e-learning 環境の開発も平行して行なっていく。

最後に、本テストと他のテスト等との関係を検証していく。質問力は創造性や問題解決能力に繋がっていると考えられるため、問うている論理の種類とこれらの能力や資質との関係を調べる。創造性については Torrance や Guilford の検査などを利用する。問題解決力については、問題解決力検定協会が行っている問題解決力検定や PICK&MIX 社の PROG のリテラシーテストなどを受けた結果との相関を検証する。PROG には周囲の状況に対応するための意思決定や行動指針などの特性を調べるコンピテンシーテストもあるため、このようなコンピテンシーとの相関についても検証する。さらに、高校生と大学生については通常の学力試験の成績との相関についても検証するとともに、探究活動などにおいてどのような行動を取っているかなど、定性的かつ質的なデータを収集し、相関を調べていく。社会人について、例えば企業の人事担当部署と連携し、QQ と人事考課との相関についても調査したい。

## 謝辞

本研究は科研費(19109437)の助成を受けたものである。

## 参考文献

- 安彦忠彦(2016)。「習得から活用・探究へ」溝上慎一編『高等学校におけるアクティブラーニング理論篇』東進堂, 70-72.
- Holmses, W., Bialik, M., Fadel, C. (2019). *Artificial Intelligence in Education*, Center for Curriculum Redesign (関口貴裕編訳(2020).『教育 AI が変える 21 世紀の学び』北大路書房).
- 池田文人(2018)。「どのように質問力を評価するか?メタモデルに基づいた評価の信頼性に関する検証」『高等教育ジャーナル: 高等教育と生涯学習』**25**, 55-61.
- Ikeda, F. (2021). "What are Scientific Questions?," *Journal of Higher Education and Lifelong Learning*, **28**, 35-43.
- 松尾知明(2015).『21 世紀型スキルとは何か コンピテンシーに基づく教育改革の国際比較』明石書店.
- 中島英博編著(2018).『シリーズ大学の教授法 学習評価』玉川大学出版部.
- Trow, M. (1961). *The Second Transformation of American*

- Secondary Education*, Oxford University Press (天野郁夫訳 (1981). 『アメリカ中等教育の構造変動』東京大学出版会) .
- Einstein, A. and Infeld, L. (1939). *The Evolution of Physics*, volume 14, Cambridge University Press (CUP).
- Young, C. (2014). “Asking the right questions,” Retrieved March 8, 2021 from *The Guardian, Newton Channel, Science* Web site:  
<https://www.theguardian.com/science/2014/jan/28/asking-right-question>
- OECD (2019). “Chapter: PISA 2018 Science Framework,” *PISA 2018 Assessment and Analytical Framework*, OECD.
- 池田文人, 山本恭裕, 高田眞吾, 中小路久美代 (1999). 「コミュニティ知識ベース環境の構築へ向けての知識の形成と利用に関する調査と分析」『情報処理学会論文誌』**40**(11), 3887－3895.
- Voltaire. (2021). “Voltaire Quotes,” Retrieved March 8, 2021 from *BrainyQuote.com* Web site:  
[https://www.brainyquote.com/quotes/voltaire\\_100338](https://www.brainyquote.com/quotes/voltaire_100338)
- Kant, I. (2007). *Critique of Pure Reason*, Penguin Books, English translated edition.
- Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*, The University of Chicago Press.
- 藤井利江・山口裕幸 (2003). 「大学生の授業中の質問行動に関する研究—学生はなぜ授業中に質問しないのか?—」『九州大学心理学研究』**4**, 135－148.