

# 大学入学共通テストへの段階表示換算表導入の論理と現状

沖 清 豪

## 1. 問題意識と課題設定

本稿は、2010年代に進められた高大接続改革議論の成果として、2021年度から大学入学共通テストで実施されている成績の段階表示をめぐる議論を検証し、改革議論において繰り返し言及されていた段階表示（段階別評価）が現在までなぜ十分活用されているとは言えない状況となっているのかについて検討するものである。

### (1) 背景

2010年代を通じて進められた高大接続改革は、本来は高校教育改革と大学教育改革を一体に進めるために、その接続としての入学者選抜も併せて改革することが目指されたものであった。しかし、特に選抜制度改革については、大学入試センター試験から大学入学共通テストへの変更など改革案の一部は実現したものの、議論の当初から提案されていた多くの改革案が検討途中で断念されるなど、多くの混乱をもたらしたと評価されている。実際に大学入試のあり方に関する検討会議の提言では多くの問題点が指摘されることとなった（大学入試のあり方に関する検討会議 2021）。

さらに、議論を通じて実現された改革についても、そのいくつかについてはすでに修正が図られたものや十分に活用されていないものがみられる。前者については調査書の記載内容についての修正を指摘できる（文部科学省 2021a）。この修正では、「調査書は指導要録に基づき作成する原則や、学校の働き方改革を受けた教員の負担軽減の観点も踏まえて、簡素化された新しい指導要録の参考様式に合わせて、例えば「7. 指導上参考となる諸事項」の欄を簡素化するなどの様式の見直しを行うとともに、枚数は表裏の両面1枚とする。」（文部科学省 2021a: 別紙2）とされ、2025年度入学者選抜から記載内容を削減する方向へと再転換している。

そして十分に活用されていないものとして、大学入学共通テストに導入された成績の段階表示が注目される。入学者選抜改革の議論においては、大学入試センター試験という全国共通テストについて1点刻みの点数を表示する現状の方式から、段階別評価を導入して結果の段階表示を主張する意見が繰り返し表明されてきた。今回の高大接続改革においては、従来の1点刻みの点数

表示は残されたものの、各受験生の得点に基づいて、スタナインと呼ばれる方法で各教科・科目の200点ないし100点満点を9段階に分けた表示が導入され、段階表示の換算表に基づいて、各教科・科目の得点に応じた段階表示を確認できるようになった。

しかしながら、導入後数年の現時点において、個別大学における入学者選抜でこの段階表示を活用している例はわずかである。実際、段階表示の導入一年目においては、入学者選抜のために、あるいはそれ以外の目的で段階表示を使用した大学はなく、導入を検討している大学もあるが導入にあたっての課題が指摘されている（文部科学省 2021b）。

また、2022年1月25日の文部科学大臣記者会見での質問において、直前に実施された大学入学共通テストの段階表示の課題が指摘されている。具体的には「合否判定に使っていただく前提で出されていると思うんですが、48点と100点が同じ評価になるようなもの（数学Ⅰのレベル9の得点範囲：引用者注）を合否判定に使える」のか、そして「得点が出ている以上、この評価（段階表示：引用者注）は要らない」のではという記者の質問に対して、大臣は懸念を示しつつ、今後の改善に期待する内容で応答している（文部科学省 2022）。しかしながら、現在に至るまで、この段階表示について制度的にも実践的にも改善が図られてはいない。

## （2）語彙の整理

ここで、語彙の整理をしておきたい。冒頭で示した「段階別評価」とは、文部科学省の資料、例えば「大学入試のあり方に関する検討会議」で毎回提示されている参考資料分冊において、「大学入学共通テストにおける段階別評価」というスライド資料で使用されている表記である。2013年以降のいくつかの会議等で段階表示がどのように議論されていたかを示す総称として使用されている。これは、共通テストで採用されたスタナインの段階表示を含め、より一般的な学力試験における段階表示をするための評価制度として捉えた場合に用いるものと考えられる。文脈に応じて「段階別表示」という表記も確認できる。

一方、「段階表示」は現在、共通テスト実施要項およびスタナイン導入に関する文書で使用されている表記であり、多くの先行研究はこちらの表記に統一されつつある。本稿では、原則として「段階表示」を用いつつ、一般論ないし個別文書で具体的に言及されている場合に「段階別評価」「段階別表示」を用いることとする。

## （3）スタナインの概要

さて、今回の大学入学共通テストで導入された成績の段階表示は、各大学のアドミッション・ポリシーに応じた受験生の多様な評価方法において活用することを目的として、従来の科目別得点に加えて、当該科目の受験生全体における各受験者の成績順位をスタナインに基づく1から9までの9段階で表示するものである。このスタナインという区分法は、学力テストで段階別評価

## 大学入学共通テストへの段階表示換算表導入の論理と現状

を行う際の根拠となる理論として、2000年代以降紹介されてきた考え方である。具体的には、もっとも成績の高い段階が9で受験生の4％程度、以下8が7％（上位から11％）、7が12％（同じく23％）、6が17％（40％）、5が20％（60％）、4が17％（77％）、3が12％（89％）、2が7％（96％）、1が最後の4％となる（大学入試センター 2023）。

実際の大学入学共通テストに採用されている段階表示は、以下の原則に従っている。

- ① 各教科・科目ごとに9段階表示を行うとともに、理科①については合計点についても9段階表示を行う。
- ② 英語はリーディングとリスニングを分けて段階表示を行う。
- ③ 個別受験生の各教科・科目の得点に基づく段階表示への換算表は、試験実施後に大学入試センターから別途発表される（大学入試センター 2023）。

実際に導入されたのち、前述したように2022年度大学入学共通テストにおける数学Ⅰと数学Ⅰ・数学Aという数学の科目間で得点と段階表示との乖離が大きくなったことが注目された。例えば数学Ⅰで48点となった受験生は段階表示では最高位の9となるのに対して、数学Ⅰ・数学Aで48点となった受験生の場合には段階表示は6となり、同一教科内の選択科目間で、かつ数学Ⅰが共通する両科目で同一の得点となったのにもかかわらず、段階表示では3段階の違いが生じているという事例が生じた。

また、2024年度大学入学共通テストにおいては、地理歴史の世界史Bと日本史Bとの間で、

表1 同一教科内科目間での得点と段階表示との乖離の事例

| 2022年度 | 数学Ⅰ    | 数学Ⅰ・数学A | 2024年度 | 世界史B   | 日本史B   |
|--------|--------|---------|--------|--------|--------|
| 段階     | 得点範囲   | 得点範囲    | 段階     | 得点範囲   | 得点範囲   |
| 9      | 48-100 | 71-100  | 9      | 94-100 | 85-100 |
| 8      | 39-47  | 60-70   | 8      | 88-93  | 78-84  |
| 7      | 31-38  | 50-59   | 7      | 79-87  | 69-77  |
| 6      | 23-30  | 41-49   | 6      | 69-78  | 61-68  |
| 5      | 16-22  | 32-40   | 5      | 54-68  | 52-60  |
| 4      | 12-15  | 24-31   | 4      | 41-53  | 43-51  |
| 3      | 8-11   | 17-23   | 3      | 31-40  | 34-42  |
| 2      | 5-7    | 11-16   | 2      | 24-30  | 27-33  |
| 1      | 0-4    | 0-10    | 1      | 0-23   | 0-26   |

例：素点48点の場合、数学Ⅰは9、数学Ⅰ・数学Aは6

例：素点78点の場合世界史Bは6、日本史Bは8

出典：令和4年度・6年度大学入学共通テスト段階表示換算表

やはり同様の乖離が生じている。例えば、世界史 B で78点となった受験生は段階表示では6となるのに対して、日本史 B で78点となった受験生の段階表示は8となり、段階表示で2つの違いが生じている。この結果は、仮に段階表示を採用した大学が、地歴のいずれかの科目で段階表示7以上を合格ラインとして設定していたとする場合には、78点となった日本史 B の受験生は合格するのに対して、世界史 B を選択した受験生は不合格になることを意味する。

こうした関係科目間での得点と段階表示とのずれが生じた場合、従来から得点での合否判定が公平・公正であると考えられてきたとされる日本の大学入学者選抜においては、このような違いが生じたことで公平性が担保されていないと感じられてしまう可能性がある。

しかし、スタナインによる段階表示の場合には当該科目の受験生が獲得した得点の分布に基づいて段階が決定されることになるので、段階の決定にあたっては公平・公正性が担保されているというのが、スタナイン導入の論理でもある。

では、こうした特性を有するスタナインやその段階表示に基づく大学入学者選抜について、これまでどのような研究が蓄積されてきたのであろうか。

#### (4) 先行研究

本稿の課題については、大きく二種類の先行研究が確認できる。

第一の先行研究は、テスト理論に基づいて、段階別評価を大規模学力テストに適用する方策に関する一連の研究である。莊島（2010）はニューラルテスト理論とスタナインの妥当性について検討しており、莊島（2018）では学力テストにおけるスタナインの利用について整理している。さらに、莊島・石岡他（2023）ではテスト理論を踏まえた分位点設定の妥当性について論じられている。

スタナインに基づく段階表示を実際の大学入学共通テストに導入する場合の考え方や妥当性については、橋本・莊島他（2021）および前川・大津（2021）が検証している。テスト理論に基づく限り、スタナインに基づく段階表示には一定の合理性があることが示されている。なお、スタナインに限定されないものの、令和3年度大学入学者選抜で学力テストなどの結果を段階表示したデータに基づいて合否判定を行っている大学の数については大学入試センター研究開発部（2023）が全国調査を行っている。

では実際に導入する場合の大学側で検討すべき点は何であろうか。林・高木（2021）は導入を踏まえた論点整理を行い、林（2022）は実際に段階表示に基づいた合否判定を行う場合に考慮すべき点を示している。さらに、森川他（2024）は2025年度から導入を予定している鳥取大学における選抜制度の設計について説明している。

以上のように、スタナインの研究やそれに基づく段階表示についての先行研究は一定の蓄積があるものの、スタナインの段階表示がなぜ現在の入学者選抜において活用されないのかについて

は、少なくとも実証レベルで検討されてきたとは言いがたい。

#### (5) 問題意識・研究仮説・研究方法

以上を踏まえて、本稿では「なぜスタナインの段階別表示は現在の入学者選抜において活用されないのか」を大きな問いとして設定し、「段階別表示に関する導入議論において、理想と現実が整理されないまま、制度だけが導入されたためではないか」という研究仮説を政策立案過程の議論に基づいて検証することで、今後の制度改革への示唆を得ることとしたい。

具体的には、何が批判されていたのか（前提）、何が目指されていたのか（目標）、誰が何を主張していたのか（理想）、想定されていた目標はスタナインの段階別表示で達成可能だったのか（現実）、および何を改善しないと理想が実現されないのか（障壁）について検証していく。

この検証にあたっては、大学入学共通テストへのスタナインの9段階別表示導入が議論された会議体の議事要旨・概要を対象とする。そして、どのような主張が展開され、どのような制度を導入することが想定されていたのかを明らかにする。

## 2. 高大接続システム改革会議における議論

本稿では、段階別評価導入までの議論を精査する対象として、二つの会議体に注目する。一つが高大接続システム改革会議であり、もう一つが本改革会議後に設置された「大学入学希望者学力評価テスト（仮称）」検討・準備グループである。もちろん、これらの会議体で検討される以前から、すなわち中央教育審議会をはじめとした入試制度改革や高大接続をテーマとして勧告・提言を行った各種審議会において、大学入試に用いられる学力テストの段階別評価に言及されてきた。本節では、システム改革会議までの議論の経緯、およびシステム改革会議における議論を確認する。

#### (1) 高大接続システム改革会議までの議論の流れ

本項では、2014年に開始された高大接続システム会議までの審議会等において、大学入試、特に全国テスト（大学入試センター試験）における段階別評価の導入に言及している答申の議論を検討する。

##### 1) 中央教育審議会（1999）における大学入試センター試験の段階別表示議論

高大接続に注目し、入試の公平性を過剰に重視する考え方からの転換を示唆した1999年の中央教育審議会答申では、試験結果の得点表示から段階別表示への転換について以下のように言及している。

各大学が受験生の多様な能力・適性等を評価し、それぞれの教育理念等にふさわしい資質を持った学生を見いだすためには、大学入試センター試験の利用方法は、得点による選抜だけでなく、大学入試センター試験の成績が一定の水準に達していれば、その後は、学力試験以外の選抜資料で合格者を決定する方法や、大学入試センター試験の成績を概括的にまとめ、それぞれに成績としては同質のグループとして扱った上で、そのグループに応じ異なった個別試験を実施する方法など、いわば資格試験的な取扱いを含め、各大学の多様な利用方法が、それぞれの創意工夫により推進されることが望ましい。

なお、大学入試センター試験の成績表示を得点ではなく段階別表示とすることについては、各大学が受験生に最低限要求する当該試験の成績は各大学によって異なり、したがって、同質のグループの定義や扱い方も各大学によって異なってくるものであるため、大学入試センター試験の成績は得点表示とした上で、各大学においてこれを自由に利用しうるようにすることが適当である。

すなわち、大学入試センター試験の得点表示を個別大学の創意工夫で活用して資格試験的に使うこともありえらるしつづ、この試験の成績表示をすべて段階別表示とすることについては消極的な姿勢が示されており、あくまで各大学において得点表示を多様な基準・尺度として用いることが想定されている。また、「一定の水準」が何を示すかについての言及はなく、結果的に各大学で合否判定を行うにあたり、任意の水準を設定することが想定されている。

## 2) 大学審議会（2000）における大学入試センター試験の段階別表示議論

中央教育審議会（1999）を踏まえつつ、より公平性に関する論点に注目していた2000年の大学審議会答申では、段階別評価について、「1点差刻みの絶対的公平性のみに固執するのではなく、それ以外の多様な試みが、大学の自主的な判断やそれぞれの入試設計に基づいて行われ、そのような取組が、社会においても許容されていくことが必要」であると主張している。また、「得点による選抜だけでなく、一定の学力水準に達しているかどうかの判定に主として用いる」べきと言及している。

一方、試験自体の「難易度の標準化に関する研究は重要」であり、「1点差刻みの絶対的公平性のみに固執するのではなく、一定の学力水準に達しているかどうかという大まかな達成度を見るという視点に立てば、資格試験的な利用は十分可能」としている。

こうした議論で繰り返し批判されているのは、大学入試における学力試験に対して、社会全体が1点差刻みの絶対的公平性を過剰に求めてきたことである。そして、大学入試センター試験については各大学における「資格試験的な」到達度試験としての活用を強調している。ただし、答



申内では段階表示の導入に言及しているわけではなく、理念的な議論に留まっている。

### 3) 中央教育審議会（2008a、2008b）での議論

2008年の中央教育審議会答申、いわゆる学士課程答申では、答申案策定段階のワーキンググループで高大接続テスト（仮称）の導入が提言されたが、答申自体では本テストの議論に慎重な姿勢が示された。

ワーキンググループの審議のまとめでは、「上記以外の「学力担保」措置の選択肢として、高校・大学が協力して AO 入試や高校の指導改善に活用できる新しい学力検査（高大接続テスト（仮称））を実施することも有効な方法である。」として、選抜性の高い現行の学力テストだけでなく、新たにより多くの高校生の基礎学力の確認や多様な生徒が学んでいる高校における学習指導の改善に活用可能な学力テストとして高大接続テスト（仮称）の新設を想定している。そのうえで、「大学入試センター試験については、多くの大学で個別学力検査との組み合わせで行われ、大学進学希望者の基礎学力を評価するものとして汎用性が高く、高校・大学関係者から厚い信頼を得る一方、科目の選択幅などについて改善の可能性が指摘されている。」（中央教育審議会 2008b）としており、高く評価しつつも改善の可能性が指摘されており、結果的に大学入試センター試験を高大接続テスト（仮称）に転換する可能性が示唆されている。

こうしたワーキンググループでの議論を踏まえつつ、答申本文では、高大接続テスト（仮称）は「学力を客観的に把握する方法の一つとして一定の意義があると考えられる一方、高等学校教育の在り方との関係上、留意すべき点も種々あることから、高等学校及び大学関係者間の十分な協議・研究が行われることを期待」（中央教育審議会 2008a）するとし、大学入試センター試験の抜本的改革は想定せず、「学力不問」の選抜における基礎学力を確認するテストとして高大接続テスト（仮称）を別途設計することを想定しつつ、その検討については協議と研究の継続を主張し、新テストへの慎重な姿勢を維持している点が注目される。

ただし、いずれの議論においても、大学入試センター試験の改革としての段階別評価の導入、あるいは高大接続テスト（仮称）に具体的に段階表示を導入するといった具体的な提案には至っておらず、原則論に留まっている。

### 4) 教育再生実行会議（2013）の議論と提言

高大接続改革を実質的に牽引する文書となった2013年の教育再生実行会議第四次答申は、現行の大学入試センター試験や個別大学の学力選抜試験を「知識偏重の1点刻み」で合否判定を行っているとし繰り返し批判し、「知識偏重の1点刻みの試験のみによる選抜や、逆に、学習への意欲や努力の減退を招くような学力不問の選抜によって、本来伸びるはずの若者の能力を損ねることがあってはなりません。」（教育再生実行会議 2013）と主張している。

その上で、本答申では到達度テスト（発展レベル）（仮称）の新設が提唱され、「その結果をレベルに応じて段階別に示すことや、各大学において多面的な入学者選抜を実施する際の基礎資格として利用することなど、知識偏重の1点刻みの選抜から脱却できるよう利用の仕方を工夫する。」（教育再生実行会議 2013）としている。

なお、本答申では、到達度テスト（発展レベル）（仮称）での段階別評価を提唱するにとどまっており、学校における指導改善や、生徒の学習意欲の喚起及び学習改善を目的として提唱されている到達度テスト（基礎レベル）（仮称）に関する議論では、結果の段階表示には言及されていない。また発展レベル（仮称）の段階別評価についてのレベルをどのようなものとして設定するかについても特段の言及はなく、あくまで理念の提示に留まっている。

#### 5) 中央教育審議会（2014）における提言

2008年の答申では強調されなかった「1点刻み」批判が中央教育審議会答申でも表出したのは、2014年の中教審答申からである。本答申では、「1点刻み」の客観性にとらわれた評価から脱し、各大学の個別選抜における多様な評価方法の導入を促進する観点から、大学及び大学入学希望者に対して、段階別表示による成績提供を行う」（中央教育審議会 2014）として、成績提供について1点刻みによる得点から段階別表示へと転換することを提言している。

一方、「段階別表示の具体的な在り方や、あわせてどのようなデータ（標準化得点や、パーセントイル値に基づき算出されたデータ等）を大学に提供することが適当かについては、別途、専門家等による検討を行うこととする。」（中央教育審議会 2014）とし、具体的な設計については、議論が十分ではなく専門的な観点での検討が必要であるとしており、実際の段階別表示への転換には慎重な姿勢を示すにとどまっている。さらに、そうした設計が可能かどうか、実現可能性に言及されているわけでもなく、依然として理念レベルの議論に留まっている。

#### (2) 高大接続システム改革会議の目的と委員構成

以上のような経緯を経て、大学入学共通テストの素案が検討される中で、文部科学省や関連組織は高大接続改革に舵を切ることになった。中央教育審議会「高大接続改革答申」が2014年12月22日に公表されたのち、「高大接続改革実行プラン」が2015年1月16日に公表され、高大接続システム改革会議が組織された。

本会議は各界代表27名の委員から構成され、「高等学校基礎学力テスト（仮称）及び大学入学希望者学力評価テスト（仮称）の在り方について」「個別選抜（各大学が個別に行う入学者選抜をいう。）の改革の推進方策について」および「多様な学習活動・学習成果の評価の在り方について」という課題について検討された。

第1回の会議が2015年3月5日に開催され、その後ほぼ一月に一回程度の会議が開催され、6



回の会議を終えた2015年9月15日に中間まとめが公表されている。さらにその後8回の会議が開催され、2016年3月31日に最終報告が公表されている。

以下、議事録に基づいて、どのような議論を通じて段階表示が制度化されていったのか、議論の中で残った、ないし消えた論点を確認することとしたい。

### 1) 第3回会議

2015年6月18日に開催された第3回会議では、検討されている2つのテストに段階表示を導入するにあたっての考え方が示されている。

高等学校基礎学力テスト（仮称）については、「生徒が成果を実感しやすくなるように、10段階以上の多段階で結果を提供していったらどうかということで、議論が今なされているところ」としており、大学入学希望者学力評価テスト（仮称）については、「大学、大学入学希望者に多段階で結果を提供するということと併せて、例えばパーセンタイル値などによって具体的なデータを大学に提供することを目指してはどうかといったことも議論」しているとしている。また、英語4技能資格試験の得点が10段階以上の点数で表示されていることにも言及しており、いずれについても相当に細かい段階設定を想定していることが窺われる。

また、段階表示に関する具体的な議論の進捗に関する質問が民間教育産業関係者および大学関係者側から出されたものの、この時点では新テストワーキンググループにおける専門家による事例紹介にとどまっており、具体的な検討までには至っていない。

一方、第3回会議においては、基礎学力テストに段階表示を導入することに対する強い疑念が示されている。例えば、到達度の確認ないし高卒時の学力の担保として基礎学力テストを用いるべきとの主張の一環として、「例えば10段階ではなくて、パスかパスでないでもいい」という形で基礎学力テストに段階表示を導入することへの疑念が示されている。また、テスト理論の専門家からは、「もし（問題が…引用者注）公開されないとしたら、1次元の10段階の一つのスコアを返す以上のフィードバックや診断評価機能はどのように確保するのか」と疑念が示され、基礎学力テストに期待されている学力診断やカリキュラム・マネジメントのための情報提供という機能は、段階表示の試験結果では有効に機能しないとの意見が出されていた。

関連する議論においても、少なくともこの時点で、基礎学力テストに段階別評価、段階表示を導入すべきという強い主張は確認できない。

### 2) 第5回会議と中間まとめにおける高等学校基礎学力テスト（仮称）の段階表示

2015年8月5日に開催された第5回会議では、資料1「高大接続システム改革会議（中間まとめ）（案）」における高等学校基礎学力テスト（仮称）の段階表示の説明が事務局から行われている。段階表示の目的は、「一定程度のきめ細かな段階を示すことで自身の学習の定着度を明確に

するとともに、以後の学習の目標になりやすく、学習の成果が実感しやすくなるよう、10段階以上の多段階で本人に結果の提供を行う。」こととされており、「「目標に準拠した評価」（いわゆる絶対評価）を行う性質のテストであり、「集団に準拠した評価」（いわゆる相対評価）を行う性質のテストではない」と性格や方法が規定されている。また、高校関係者との意見交換を踏まえて、適切な段階表示方法を検証していくことも言及されている。

こうした説明に対して、第3回会議と同様に、根強い批判が加えられている。

具体的には、本テスト自体が多大な（人的、金銭的）コストが必要なにもかかわらず、「どこに焦点を当てたテストなのかがはっきりしない」と基本的な目的が不明確であることを批判し、「中間まとめに、そのはっきりしない状態で何か断定的なことを書くことは避けるべきではないか」と思います。ここまで出てきた複数受検とか、段階表示とか、CBTとか、IRTとか、これらは全て本質的でない話で、なぜ複数なのか、何のために複数なのか、何のためにIRT、CBTなのか、段階表示なのか。目的がはっきりしないので、どれも定まらない」との主張が展開された。つまり、本会議が何をやるかの議論に特化し、なぜ導入するのかの議論が回避されていることが批判されており、その一例として段階表示の議論がなぜ導入するのかの議論を回避したまま進められていることへの疑念が示されている。

しかしながら、2015年9月15日に公表された中間まとめ案では、第5回会議の事務局側の説明に沿った内容の文案が公表されることとなった。具体的には、以下のようにここまでの議論がまとめられている。

一定程度のきめ細かな段階を示すことで自身の学習の定着度を明確にするとともに、以後の学習の目標になりやすく、学習の成果が実感しやすくなるよう、10段階以上の多段階で本人に結果の提供を行う。また、より詳細な学習改善や指導改善にも生かすことができるようにするため、単元ごとなど分野別の結果や各設問の出題のねらい等を提供する。その際には、適切な段階表示の在り方について検証を通じて高等学校関係者等の意見等も踏まえながら、試行を進めていくことが必要

各生徒の高等学校段階における基礎的な学習の達成状況について確認する「目標に準拠した評価」（いわゆる絶対評価）を行う性質のテストであり、「集団に準拠した評価」（いわゆる相対評価）を行う性質のテストではないことから、各生徒等の順位は示さない

### 3) 第5回会議と中間まとめにおける大学入学希望者学力評価テスト（仮称）の段階表示

第5回会議では、大学入学希望者学力評価テスト（仮称）の段階表示についても事務局から説明されている。具体的には、「結果の表示については、個別大学の入学者選抜における多面的・

総合的な評価を促進する観点から、大学や大学入学希望者に対し、結果の多段階による表示による提供を行うこと、あわせて、種々の具体的なデータ（例えば、パーセンタイル値に基づき算出されたデータ、標準化得点、出題分野ごとの正答数や誤答数など）を大学に提供することなどについて、大規模な共通テストとしての幅広い識別力の確保の必要性なども踏まえつつ、今後より専門的に検討する。」とされた。つまり、段階表示を前提としつつ、追加的に提供するデータを含めて専門的な検討に委ねるとして、断定的に結論や方向性を示すことが回避されている。

結果的に中間まとめでは、第5回会議での説明にほぼ沿った形で、

個別大学の入学者選抜における多面的・総合的な評価を促進する観点から、大学や大学入学希望者に対し、結果の多段階による表示による提供を行うこと、あわせて、種々の具体的なデータ（例えば、パーセンタイル値に基づき算出されたデータ、標準化得点、出題分野ごとの正答数や誤答数など）を大学に提供することなどについて、大規模な共通テストとしての幅広い識別力の確保の必要性なども踏まえつつ、今後より専門的に検討する

と整理されている。結果の多段階表示を前提としつつ、他の様々なデータの提供も求め、今後の専門的な検討を要件としたのである。

#### 4) 中間まとめ後の議論

中間まとめが公表されたのち、本会議の中では段階表示に言及した事例は2つに留まる。

第一の事例として、第7回会議で2つの試験の統合をめぐるヒアリング資料の中で、PTA 連合会会長が、二つのテストを実施するのではなく、一つのテストを多数の段階を設定した段階別評価を実施することで、基礎学力と大学入学者の学力を確認することができるといふ論旨において段階別評価に言及している。しかし、こうした主張について、中間まとめ公表以降において、段階表示についての具体的な検討が行われたとは言い難い。また、結果的に本会議の最終報告書までに2つのテストの統合が議論されることはなかった。

第二の事例としては、第11回会議（2016年2月17日開催）と第12回会議（2016年2月24日開催）において配布された資料1「「高等学校基礎学力テスト（仮称）」の位置付け等の更なる明確化」において、基礎学力テスト（仮称）を「高校生の多様性を踏まえ、同一問題・一斉実施の方式ではなく、複数レベルの問題から学校が選んで受検し、生徒の基礎学力の定着度合いに応じた目標標準型の評価を段階表示で結果提供することや、学校行事等を考慮しながら実施時期を柔軟に設定できる仕組みとする。」と説明している。

結果的にその後の会議でも最終報告書案の表現の微修正が中心で、ここで示された論点の変更されなかった。また、大学入学希望者学力評価テスト（仮称）の段階表示については議論がなさ

れていないまま、最終報告書が公表されることとなった。

本会議の最終報告は、2016年3月31日公表された。本最終報告では、高等学校基礎学力テスト（仮称）に関しては、「高校生の多様性を踏まえ、同一問題・一斉実施の方式ではなく、複数レベルの問題から学校や受検者が選んで受検し、生徒の基礎学力の定着度合いに応じた評価を段階表示で結果提供すること」（p.21）、および「一定程度のきめ細かな段階を示すことで自身の学習の定着度合いを明確にする」（p.27）ことを目的とするものであると設定された。

また、大学入学希望者学力評価テスト（仮称）に関しては、「結果の表示については、従来の合計点方式のみでは得られない、よりきめ細かい評価情報により、個別大学の入学者選抜における多面的・総合的な評価を促進するため、多様な情報（例えば、得点だけでなく、各科目の領域ごと、問いごとの解答状況も合わせて提供するなど）を各大学に提供する。」（p.55）という表現となった。

ここで段階表示に関する議論の変化を整理すると、高等学校基礎学力テスト（仮称）については、当初の議論では10段階以上での段階表示が想定されていたのに対して、最終報告では、多段階の段階表示は維持するとしたものの、導入目的が学力診断やカリキュラム・マネジメントであることを強調するなど、当初の理念とは異なるものとなっている。

また、大学入学希望者学力評価テスト（仮称）については、段階表示の導入について会議内で強い疑義が表明された結果、最終報告では段階表示に言及されていない。つまり、段階表示を導入する理念の議論が回避される形となったのである。

### 3. 「大学入学希望者学力評価テスト（仮称）」検討・準備グループにおける議論

本節で検討するのは「大学入学希望者学力評価テスト（仮称）」検討・準備グループ（以下、検討・準備グループと略称）において、段階表示がどのように議論されていたのかについてである。

検討・準備グループは、2016年5月28日の文部科学省プレスリリース「高大接続改革の検討・推進体制について」で設置が公表された文部科学省改革推進本部・高大接続改革チームの下におかれた4つの検討組織のうちの一つである。

検討・準備グループは、高校側委員4名（現元校長各2名）、大学側委員4名（国立大学協会、公立大学協会、私立大学連盟、および私立大学協会）、テスト理論の研究者1名の計9名からなる委員で構成された。また、第4回の会議から大学入試センター内に設置された「新テスト実施企画委員会」（第1回）と合同開催の形がとられた。この新テスト実施企画委員会は入試・テストの研究者と英語教育の研究者が参加している。検討・準備グループとしては、12回の会議が開催され、その中でも段階表示に関する議論が第3回、第7回、第8回、第10回および第11回と繰り返し実施されている。以下、それぞれの会議でどのような議論があったのかを確認する。

### 1) 第3回会議と第7回会議

第1回会議では英語の成績表示について、また第2回会議では米国のSATについて紹介されたものの、日本で、特に共通テストで段階表示をどのように導入するかについては言及されていない。段階表示について本格的に議論が行われたのは第3回会議（2016年8月23日）からである。

第3回会議では、元高等学校長の委員から、「私は、共通テストは選抜でないと思っている。資格試験的になるべき」との発言に続いて、「今回の改革で言うと、もともと段階別ということがあった。細かな点数とならないようにということで始まっていたと思う。Aができればいいよとかにしていくべき。だから複数回ができる。でないと1回でやるべき。なぜ複数回を出したかを考えると、段階別で出していく。」からであると説明しており、共通テストの複数回実施と段階別評価の導入による共通テストの資格試験化が主たる目標であると主張している。一方で、他の委員からは「選抜でなくなると、基礎学力テストに近くなる。今の選抜の形式で、複数回やったら、良いことはない」との批判もみられた。つまり、共通テストが現状では「選抜試験」であることを前提としつつ、今後「資格試験」と位置付けるべきか「選抜試験」のままで改革を進めるべきかで委員間での意見の相違がみられ、この対立と連動する形で段階別評価が議論されているのである。

この関係への言及は第7回会議でも確認できる。テスト理論の研究者である委員からは、「試験は資格試験と選抜試験の2通り種類がある。資格の場合は、ここまでの技能を持っている者は全て合格となる。一方、大学でやっている選抜入試は上から100人とり、ということで、そこが大きく違う。単なるパーセントでは切れない」、「バンド幅では1点違いが、その後の選抜に効いてくる」という二つの試験の特性が説明され、選抜試験として実施する場合に、バンド幅がたとえ1点であったとしても選抜の結果が異なる可能性があること、資格試験の場合は一定水準を超えていると判断されれば合格者数は多くなる可能性があること等に言及しており、「選抜試験と資格試験の絡みを、どのように折り合いつけた形で議論するか」が重要であると整理されている。

しかしその後、二種類の試験について本会議で継続的に検討されたわけではなく、この論点は残されたままとなった。

### 2) 第8回会議以降の議論

2017年2月21日に開催された第8回会議では、原則論に戻る形で段階表示の推進を主張する委員から、「1点刻みからの脱却をしてほしい。真の学力をつけるという意味で、段階別の評価はしていかなければならないと思う。子どもにどういう力をつけさせたら良いのかということで、点数が高ければ良いということではないというのが今回の改革のキモだと思う。」という発言があり、改めて1点刻みの選抜試験が批判されている。

また、本会議では、センター試験と一般選抜の英語の成績との相関をみるための得点分布、20



点刻みで区分した分布、および受験者を得点順に並べて10%ないし20%刻みで区分した場合の分布や相関係数を研究した成果が紹介された（資料3）。その結論として、いずれの区分でも相関係数に大きな違いが生じないことが示され、「どうしても1点刻みでないと選抜できないものなのか」という疑問に対して、段階表示でも学力がかなり正確に測定されていることが示されるとともに、段階表示については「段階別表示を換算して個別試験の点数に加算すること、達成度として使用すること」といった多様な選抜基準として活用可能であることが説明されている。

高大接続改革の議論の混乱を経て、従来非公開であった会議が公開された第10回会議以降でも、毎回原則論に回帰する形で選抜試験において1点刻みの評価が実施されていることに批判を述べる委員がおり、第11回会議では「これまで以上に1点刻み選抜に利用されるといったような誤った活用がなされないよう留意する必要」、第12回会議では「もともと入試の選抜を1点刻みで行うことに大きな弊害があったので、得点の提供は、いずれはやめてほしい」との発言が確認できる。

### 3) 議論の整理

検討・準備グループの会議では、本来実際の試験のあり方に関して具体的提案を検討するはずであったところ、高大接続改革の性格・目標をめぐる理想と現実の議論が交錯し、繰り返し改革のあるべき姿が語られるという状況が生じていた。共通テストの性格については、本テストを引き続き選抜試験と位置付けるか、資格試験として再定義するかという論点が提示されたものの、議論が深まることはなく、共通テストの性格が再定義されなかった。一方で成績を段階別表示するのが一般的な英語四技能試験や記述式出題という改革が頓挫する中で、段階表示の具体的なイメージが丁寧に議論されぬまま、ただ2020年度大学入学共通テストにおいて、スタナイン導入による段階表示が事実上決定していたということになる。

### 4) スタナイン導入決定まで

検討・準備グループの解散後、スタナイン導入に関しては、大学入試センター「新テスト実施企画委員会」、およびその継承である「大学入学共通テスト企画委員会」に議論の場が移ることとなった。

実際には、2018年4月25日（第1回）から同年9月19日（第20回）まで開催された「新テスト実施企画委員会」では、段階表示に関する議事を確認することができない。その後再編された「大学入学共通テスト企画委員会」では、2019年12月16日の第8回会議の議題2「大学入学共通テスト企画委員会マーク式問題段階表示検討WGの報告について」において、スタナインの概要について説明があり、翌2020年10月2日に開催された第14回会議の議題3「令和3年度大学入学者選抜に係る大学入学共通テストにおける段階表示の取扱いについて」で原案に関する議論の結果、

スタナインによる段階表示の導入が承認されている。なお、この時点で表示の科学的な原理は説明されたが、活用する方法に関する議論は確認することができない。結果的に活用方法の例示なきまま制度導入が先行した結果となったのである。

#### 5) スタナイン導入後の混乱

2021年度大学入学共通テストから導入されたスタナインによる9段階表示は、本稿の冒頭で言及したとおり、2022年度選抜において数学Ⅰと数学Ⅰ・数学Aとの間で、同一得点であるにもかかわらず段階表示が3段階も異なる結果となり、スタナインの特性を十分理解しない限り、表示そのものに疑義が生じうる事態となった。また報道こそされなかったが、2024年度選抜においても地理歴史の科目である世界史Bと日本史Bとの間でも、同一得点であっても段階表示が2段階異なるという事例も確認されている。

受験者層が異なる教科・科目間で同一得点でありながら段階が異なるという現象は、スタナインの特性上問題のない結果であるにもかかわらず、従来からの公平観に基づく直感的に疑問を感じさせる結果であるともいえる。こうしたスタナインの性格自体が十分理解・共有されていないと思われる混乱が大学入学者選抜協議会でも生じている。

2021年5月26日に第1回大学入学者選抜協議会が開催された折、初年度にあたる2021年度大学入学共通テストの結果に基づいて公表されたスタナインによる段階表示を活用した大学がなかったことを示した「大学入学共通テスト（スタナイン）に関する調査結果について」が参考資料6として共有され、その結果を踏まえて「令和4年度大学入学者選抜実施要項の見直しの方向性（案）」（資料6）が示された。資料6では、事務局が作成した大学入学者選抜実施要項案においてスタナインの活用を促進する一文、「段階表示の利用（追加） 共通テストの得点の段階表示（スタナイン）は、それを利用して一定の学力水準に達していることの確認をした後は、それ以外の方法と組み合わせて選抜するような、共通テストの成績の基礎資格的な利用方法等を例示。」という一文が含まれていた。

この提案に対して、参考資料6で紹介されている課題が解決していないのに、段階表示を積極的に活用することを求めることに対する批判が出されるとともに、大学入試センター長（当時）から「スタナインで一定の学力水準に達していることの確認はできません。」との発言があり、スタナインの特性自体が、資料6作成者、すなわち政策立案者間でも十分理解されていないことが露呈する事態となった。

結果的にその後の大学入学者選抜実施要項では、段階表示の利用促進の記述は追加されておらず、活用状況が改善しているとも言い難い状況となっている。

## 4. 論点の整理

### (1) 理想と現実の齟齬

本稿では2つの会議体において学力テスト結果の段階表示をめぐる段階表示がどのように議論されてきたかを確認してきた。本節では、改めて議論の背景にある「理想と現実」の齟齬をめぐる生じている葛藤について検討する。

第一の葛藤として、「1点刻み」の選抜に対する批判を上げることができる。この種の批判は、1点刻みで結果が示される学力試験における過度な公平性追求が前提としてあり、さらに相対評価に基づいた合否判定による高校生・高校教育への悪影響を指摘する形で批判されてきた。

こうした動向は2010年代において、大学入試センター試験が「1点刻み」の元凶として改めて批判され、改革の焦点となる一方、高校教育の改善のために基礎レベルの学力を段階表示することによる、学力診断やカリキュラム・マネジメントの改善という期待も語られている。しかしながら、実際には、技術的な論点を検討するはずの会議体に「現実」としての制度設計が丸投げされた結果、「理想」としての理念が十分練られないまま、技術的な議論に移行してしまい、関係者内ですら十分な理解がないまま、制度化されてしまったように思われる。

第二の葛藤として、大学入学共通テストは従来からの選抜試験のままでよいのか、新たに資格試験化すべきなのかについても、会議体において理想と現実とを明確に整理し、制度設計に落とし込む議論がなされないままであった点が挙げられる。特にスタナインによる9段階表示をどのように活用するかについては、関係者間での議論が不十分なまま制度化が進められ、結果的に活用方法が現在まで各大学に丸投げされた格好となった。このことが、導入が進まない一因となっているように思われる。

なお高大接続改革を検証した「大学入試のあり方に関する検討会議 提言」(2021年7月)では、CEFRとの関連で段階表示について言及している箇所があるものの、提言本文にスタナインに関する記載はない。このこともスタナインの9段階表示に対する社会的関心の低さを示す事例と言えるだろう。

### (2) 議論の整理

本稿の結びとして、課題設定で言及した、何が批判されていたのか(前提)、何が目指されていたのか(目標)、誰が何を主張していたのか(理想)、想定されていた目標はスタナインの段階表示で達成可能だったのか(現実)、および何を改善しないと理想が実現されないのか(障壁)について確認する。

はじめに、前提(何が批判されていたのか)については、2000年代以前から学力試験における「1点刻み」の成績表示とそれが高校教育・生徒に与える悪影響が批判されてきた点であること

が確認された。この批判は1990年代末の審議会答申でも言及されているが、その時点では段階表示自体の導入は否定的であった点も重要であろう。

では何が目指されていたのか（目標）。「1点刻み」試験の象徴としての大学入試センター試験の抜本的改革を求める声がある一方、特に基礎学力を検証する試験を導入し、段階表示によって生徒の学習進捗状況を測定することで高校教育の改善を求める声もあった。こうした議論を踏まえつつ、少子化や大学数の増加という社会変化の中で、段階表示による学力の大まかな確認、あるいは大学入学者の資格試験としての活用と、その他の多様な選抜制度による個別大学のAP（アドミッション・ポリシー）に即した選抜制度の開発・導入が求められたのである。

しかし、このように想定されていた目標はスタナインの段階表示で達成可能だったのであろうか。そもそもスタナインは受験者層の上位から一定の割合で区分するという基準設定の特性を有していることから、相対評価とならざるを得ない。また、段階設定の方法、試験内容や受験者の層（成績分布や平均点、科目間の差異）等の影響を受けるため、スタナインで一定の学力水準に達していることの確認はできない。つまり、大学入学共通テストの結果をスタナインで段階表示したとしても、そのことによって本テストが資格試験として利用することはできないのである。

結果的に、スタナインの導入は理念・目標と現実とが大きく齟齬を生み出すこととなった。それ自体が受験生に悪影響を及ぼすことはないものの、入学者選抜制度の改革はいかに実現性とその理念との合致とを丁寧に議論する必要があるのかを示す一例となったのである。

### （3）残された課題

初等・中等教育段階の教育が大きく変容し、主体的で深い学びが求められている中で、学力選抜試験の資格試験化と入学者選抜自体の改革を組み合わせることは、選抜制度でつながっている高校教育改革と大学教育改革とを促進するためにも必要なことである。特に、定員充足に課題が残る大学における選抜制度・接続・移行研究と、それに基づく継続的な大学教育改革は不可避であり、その改革の一環として何らかの形で高校教育の到達度を確認するための方策は引き続き検討していく必要があるだろう。

その際には、本稿で対象とした議論の内容をふまえつつ、理念の重視と合わせて現実との整合性、具体的な実現可能性を踏まえた検討が期待される。

また、2024年度選抜においては、静岡理工科大がスタナインによる段階評価を合否判定に活用している。さらに、2025年度選抜からは徳島大と鳥取大が、2026年度選抜からは千葉大が、それぞれ一部の学部・学科ではあるが、スタナインによる9段階評価をそれぞれ独自の方法で利用することになっており、わずかではあるが段階表示を活用する機運も生じている。

こうした中で、改めて改革当初の理念である「1点刻み」批判の正当性や資格試験としての活用のあり方について、研究と実践の両面での検討を続けていくことが重要ではないかと思われる。

## 参考文献

- (審議会答申については文部科学省ウェブサイト掲載のものを参照している。2024年9月20日最終閲覧)
- 教育再生実行会議 (2013)「高等学校教育と大学教育との接続・大学入学者選抜の在り方について」(第四次答申)
- 高大接続システム改革会議 (2018)「議事録」  
[https://www.mext.go.jp/b\\_menu/shingi/chousa/koutou/064/index.htm](https://www.mext.go.jp/b_menu/shingi/chousa/koutou/064/index.htm)
- 荘島宏二郎 (2010)「ニューラルテスト理論—学力を段階評価するための潜在ランク理論—」、植野真臣・荘島宏二郎『学習評価の新潮流』、朝倉書店。
- 荘島宏二郎 (2018)「学力テスト」日本児童研究所『児童心理学の進歩』57、pp.215-234。
- 荘島宏二郎・石岡恒憲他 (2023)「離散得点分布の分位点の推定誤差：スタナイン分位点に着目して」『大学入試研究ジャーナル』(33)、pp.106-112。
- 大学審議会 (2000)「大学入試の改善について」(答申)
- 「大学入学希望者学力評価テスト(仮称)」検討・準備グループ (2018)「議事要旨」  
[https://www.mext.go.jp/b\\_menu/shingi/chousa/koutou/083/index.htm](https://www.mext.go.jp/b_menu/shingi/chousa/koutou/083/index.htm)
- 大学入試センター (2023)『令和6年度大学入学者選抜に係る大学入学共通テスト受験案内』
- 大学入試センター研究開発部 (2023)『令和3年度大学入学者選抜における選抜資料の利用状況に関する実態調査』大学入試センター。
- 大学入試のあり方に関する検討会議 (2021)「大学入試のあり方に関する検討会議 提言」  
[https://www.mext.go.jp/b\\_menu/shingi/chousa/koutou/103/toushin/mext\\_00862.html](https://www.mext.go.jp/b_menu/shingi/chousa/koutou/103/toushin/mext_00862.html)
- 中央教育審議会 (1999)「初等中等教育と高等教育との接続の改善について」(答申)
- 中央教育審議会 (2008a)「学士課程教育の構築に向けて」(答申)
- 中央教育審議会 (2008b)「学士課程教育の在り方に関する小委員会高等学校と大学との接続に関するワーキンググループ(WG)議論のまとめ」
- 中央教育審議会 (2014)「新しい時代にふさわしい高大接続の実現に向けた高等学校教育、大学教育、大学入学者選抜の一体的改革について」(答申)
- 橋本貴充・荘島宏二郎他 (2021)「スタナインを得点調整に使うと」『日本行動計量学会大会抄録集』(49)、pp.114-117。
- 林篤裕 (2022)「段階表示を用いた合否判定方法の試行」『大学入試研究ジャーナル』(32)、pp.43-48。
- 林篤裕・高木繁 (2021)「大学入学共通テストにおける段階表示の提供開始に備えて」『大学入試研究ジャーナル』(31)、pp.239-244。
- 前川真一・大津起夫 (2021)「テスト得点の段階表示に伴う「情報の減少」について」『日本テスト学会誌』16(1)、pp.46-68。
- 森川修他 (2024)「大学入学共通テストの段階表示を合否判定に用いた新しい入試制度の設計」『大学入試研究ジャーナル』(34)、pp.163-168。
- 文部科学省 (2021a)「令和7年度大学入学者選抜に係る大学入学共通テスト実施大綱の予告」及び「令和7年度大学入学者選抜実施要項の見直しに係る予告」について(文部科学省高等教育局長通知)
- 文部科学省 (2021b)「大学入学共通テスト(スタナイン)に関する調査結果について」(第1回入学者選抜協議会配布資料)
- 文部科学省 (2022)「末松信介文部科学大臣記者会見録(令和4年1月25日)」  
[https://www.mext.go.jp/b\\_menu/daijin/detail/mext\\_00233.html](https://www.mext.go.jp/b_menu/daijin/detail/mext_00233.html)

※ 本稿は、日本学術振興会科学研究費23K02118の研究成果の一部である。