

国際バカロレアの評価における公平性の確保

江幡, 知佳
大学入試センター : 助教

<https://doi.org/10.15017/7429600>

出版情報 : 九州大学教育社会学研究集録. 30, pp.1-14, 2025-09-15. Seminar of Educational Planning, Measurement, Evaluation, Department of Education, Graduate School of Human-Environment Studies, Kyushu University

バージョン :

権利関係 :



国際バカロレアの評価における公平性の確保

Ensuring Fairness in the Assessments of the International Baccalaureate Diploma Programme

江幡 知佳

1. はじめに

国際バカロレア・ディプロマプログラム（International Baccalaureate Diploma Programme）⁽¹⁾は、16～19歳の生徒を対象とした2年間の教育プログラムである。生徒は、IB認定校⁽²⁾において所定のプログラムを履修し、一定以上の成績を収めることにより、IBディプロマ（IB Diploma）と呼ばれる後期中等教育修了資格を取得できる。この資格は、世界各国の大学への出願、入学において広く活用されている。その背景には、IBが、厳格な学業的水準を備えていると評されていることがある（Hill & Saxton 2014）。ゆえに、IBディプロマを入学者の決定のために活用する大学数は増加してきた⁽³⁾。

その一方で、IBディプロマの資格としての信頼性の根幹となる、その評価システムに対しては、懐疑的な見解も存在している。例えば、Bunnell（2011）は以下のように述べている。IBはもともと、インターナショナルスクールで学び、国際的な移動を伴う生徒の大学進学問題を解消するため、1960年代に開発された。しかし、1980年代以降、北米を中心に、インターナショナルスクール以

外の公立学校等でも導入・実施されるようになった。この動向に伴い、IBディプロマの授与に関わる試験の受験者数は爆発的に増加してきた。こうした受験者数の増加は、とりわけIBのように記述式問題を中心とした試験を実施する場合、その採点の質や一貫性を脅かすことになる（Bunell 2011）。同様に、Tan（n.d.）も、IBの急速な拡大は、IB機構による有資格採点者（qualified examiners）の確保を困難にし、このことは評価の信頼性に影響を及ぼしうると指摘している（Tan n.d.）。これらの先行研究は、IBが（図1および図2に示したような）受験者数の量的拡大と評価の質保証との両立という、構造的なジレンマに直面していることを示唆している。

本稿の目的は、先行研究が示唆している上記のジレンマを念頭に置いて、現在、IBにおいていかに評価の公平性（fairness）の確保が試みられているのか、およびその評価はいかなる課題を内包しているのかを明らかにすることとする。この目的を達成するために、IB機構が公表している評価関連のガイドライン（2節で詳述）を主な分析対象として、以下3つの問いを考察することを研究課題とする。その課題とは、第一に、IBにおいて、いか

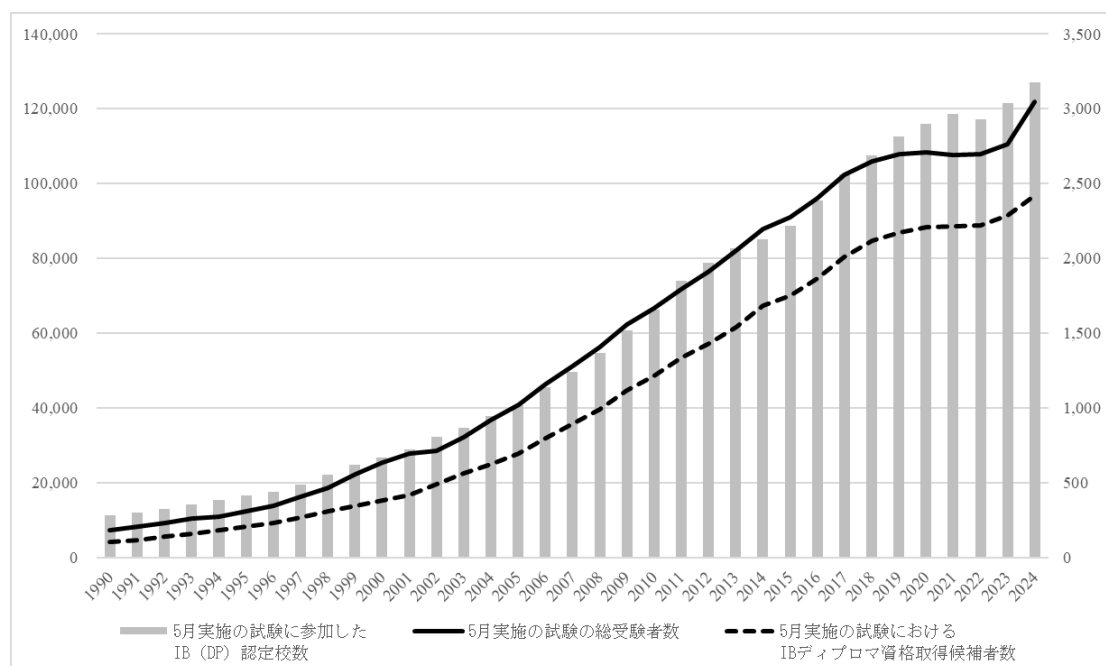


図1 IBの5月実施の試験参加校数／参加者数の推移

出典：International Baccalaureate（2022a: 11；2024: 11）をもとに筆者作成

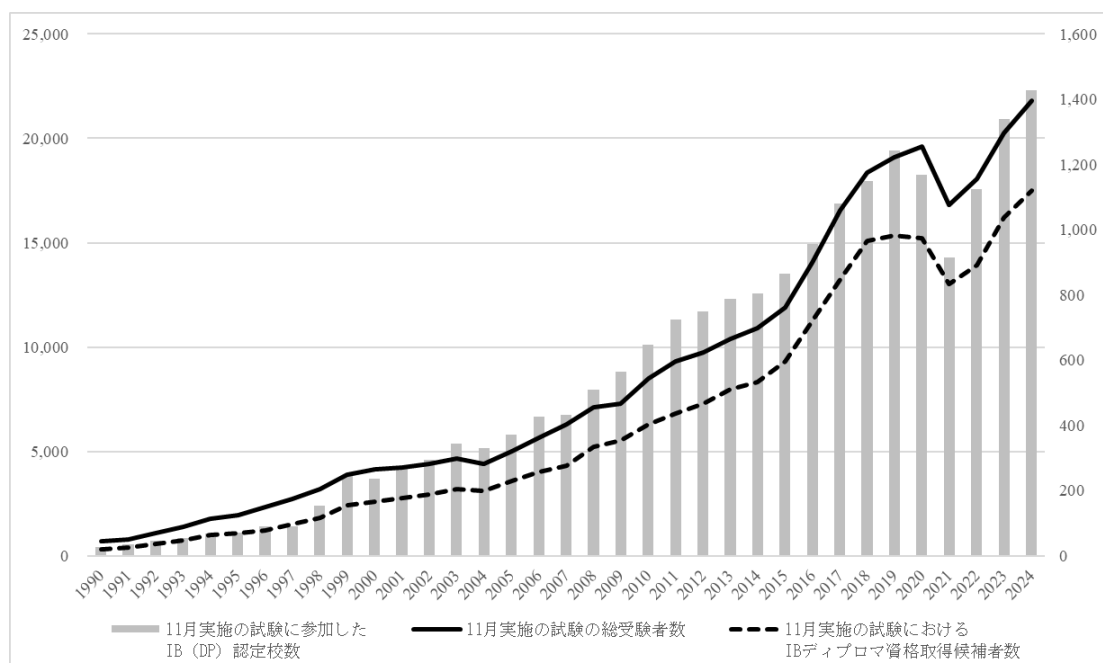


図2 IBの11月実施の試験参加校数／参加者数の推移

出典：International Baccalaureate (2023a: 11；2025: 11) をもとに筆者作成

に外部評価 (External Assessment) の質の保証が試みられているのか、第二に、IB において、いかに内部評価 (Internal Assessment) の質の保証が試みられているのか、そして第三に、評価における公平性をめぐって、現在 IB 機構の内部でいかなる課題や論点が認識されているのか、である。なお、IB ディプロマの授与に関わる最終的な成績評価は、IB 機構による外部評価と、各 IB 認定校の教員による内部評価から構成されている。

本稿の論点は以下のとおりである。IB における外部評価／内部評価は、外部試験官の階層的構成やモデレーションという仕組みの運用によってシステムチックに実施されており、ゆえに各国・各機関から一定の信頼を得ている。ただし、IB 機構も自覚しているとおり、それらは完全なものとはいえない。なぜなら、試験が国際的に実施されることに特有の課題や生成 AI の台頭に伴う課題などに、IB 機構は取り組んでいる最中だからである。換言すると、IB 機構の定義する公平な試験 (= 受験者全員に同等の機会を保障し、合理性のない有利／不利が排除されている試験) ⁽⁴⁾ の実現に向けた試行錯誤は現在も続いているためである。

国内の先行研究を概観すると、教科教育学や教育評価論の領域において、IB のカリキュラムや評価方法に関する知見が一定程度蓄積されている (例えば、中村・小笹 2022；有馬 2022)。しかし、IB の評価システムの全体を射程に含めて、いかに公平性の確保が試みられているかを紹介、検討した研究は十分でない。したがって本稿は、IB の評価システムに関わる基礎的研究として位置づけ

られる。

本稿の構成は以下のとおりである。第2節では、本稿の分析素材である IB 機構が公表している評価のガイドラインの概要を紹介する。第3節では、IB カリキュラムの概要と、IB ディプロマの授与に関わる評価の全体像を概説する。続く第4節、第5節では、それぞれ外部評価／内部評価に焦点を当てて、質保証の方策を検討する。第6節では、受験者が国際的に多様であることに由来する課題、および生成 AI の台頭という現代的状況に特有の課題を整理したうえで、IB の評価の公平性に対して若干の問題提起をする。第7節では、本稿の議論を総括し、今後の課題を指摘する。

2. 分析の素材

IB 機構は、その教育理念や教育実践に加えて、各 IB 認定校に対して課す規則や各教科の内容・方法等について、多岐にわたるガイドラインを公開している。主要なガイドラインとそれらの相互関係は、図3 (次頁) に示される。図3のうち、本稿が扱う評価関連のガイドラインは、①Assessment principles and practices、および②Diploma Programme Assessment procedures である。

前者のガイドラインについては、現在、以下の5つの文書が公開されている⁽⁵⁾。

- (1) Assessment principles and practices—Quality assessments in a digital age. (International Baccalaureate 2017a)

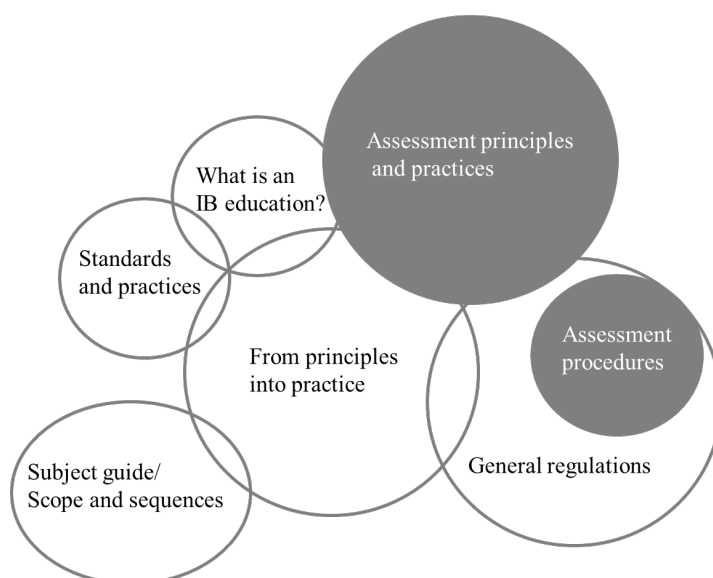


図3 IBの主要文書間の関係性

出典：International Baccalaureate(2017a: 25)

- (2) Assessment principles and practices—A guide to assessment for students and parents/guardians. (International Baccalaureate 2020a)
- (3) Assessment principles and practices—A guide to assessment for teachers and coordinators. (International Baccalaureate 2020b)
- (4) Assessment principles and practices—A guide to assessment for examiners. (International Baccalaureate 2020a)
- (5) Assessment principles and practices—A guide to assessment for chief examiners and principal examiners. (International Baccalaureate 2020d)

(1)は、IB の評価の哲学を体系的に提示した基盤的な文書である。同書は、「私たち [IB 機構] が行う評価が有意義で、公平で (fair)、関係する生徒の最善の利益になることを確認するために、IB が採用している原則を説明することを目的として」(International Baccalaureate 2017a: 5)⁶⁾いと述べ、主に外部総括評価 (external IB summative assessment/formal testing) に焦点を当てている。(2)～(5)は、それぞれ生徒と保護者、教員とコーディネーター⁷⁾、採点官 (examiner)、チーフ採点官 (Chief Examiner) と主任採点官 (Principal Examiner) を対象に、IB 機構による評価の質保証について概説する短い文書である。

後者のガイドラインとして、本稿は、Diploma Programme Assessment procedures 2023 を参照する (International Baccalaureate 2023b)。同書は、主に 2023 年 5 月 (北半球の IB 認定校の生徒が主に受験) および同年 11 月 (南半球の IB 認定校の生徒が主に受験) 実施の評価セッションで遵守すべき手続きを詳細に示すマニュアル

ルである。同様の文書が、毎年 IB 機構によって発行されている。

本稿の第 4 節以降では、主として Assessment principles and practices—Quality assessments in a digital age (International Baccalaureate 2017a) に依拠しつつ、適宜他の文書を参照しながら、前節で提示した 3 つの研究課題を検討する。

3. IB のカリキュラムと評価の全体像

本節では、IB 機構のウェブサイト⁸⁾等を参考にして、IB のカリキュラムと評価の全体像を概説する。

IB のカリキュラムは、6 つの教科グループ (subject groups) および 3 つのコア要素 (core elements) から構成される。6 つの教科グループとは、言語と文学 (Studies in language and literature)、言語習得 (Language acquisition)、個人と社会 (Individuals and societies)、理科 (Sciences)、数学 (Mathematics)、および芸術 (Arts) である。各教科には複数の科目が含まれており、生徒は、各教科グループのなかから 1 科目を選択、履修する。その際、6 科目のうち 3～4 科目を上級レベル (Higher Level) で、残りの科目を標準レベル (Standard Level) で履修することが求められる。なお、芸術の代わりに、理科、個人と社会、あるいは言語関連の科目を履修することが認められている。

IB では、生徒の幅広い教育経験や獲得した知識・スキルの応用が重視されるため、教科学習のみならず、3 つのコア要素も修了要件として設定されている。3 つのコア要素とは、知の理論 (Theory of knowledge: TOK)、課題論文 (The extended essay: EE)、および創造性・活動・奉

表1 IBにおけるコア要素への加点点表						
	知の理論					
課題論文		A	B	C	D	E
	A	3	3	2	2	資格授与なし
	B	3	2	2	1	
	C	2	2	1	0	
	D	2	1	0	0	
	E	資格授与なし				
出典：International Baccalaureate（2017a: 220）の表を筆者訳出						

表2 IBにおける内部評価と外部評価の割合				
教科グループ	科目例	レベル	内部評価の割合（％）	外部評価の割合（％）
言語と文学	言語A：文学 (Language A: literature)	標準	30	70
		上級	20	80
言語習得	古典言語 (Classical languages)	標準	30	70
		上級	20	80
個人と社会	ビジネスマネジメント (Business management)	標準	30	70
		上級	20	80
理科	生物 (Biology)	標準	20	80
		上級	20	80
数学	数学：分析とアプローチ (Mathematics: analysis and approaches)	標準	20	80
		上級	20	80
芸術	ダンス (Dance)	標準	40	60
		上級	40	60

出典：International Baccalaureate, "DP curriculum" <https://www.ibo.org/programmes/diploma-programme/curriculum/> (accessed 2025-6-25) から閲覧できる各科目のsubject briefをもとに筆者作成

仕 (Creativity, activity, service: CAS) である。TOK において、生徒は、さまざまな知のための方法 (ways of knowing) と多様な知識の形態について思慮深く目的をもって探究するため、「どのようにして私たちは知るのか」 (How do we know?) といった問いについて議論する。結果として、自らの個人的・思想的な前提をより深く認識し、文化的な視点の多様性と豊かさを理解できるようになることが期待される。TOK の評価は、世界における TOK の現れを 3 つの対象物で探究する展示 (exhibition)、ならびに TOK における概念的課題に焦点を当てたエッセイ (1,600 語) で構成される。EE とは、履修した 6 科目のうちのいずれかに関連する、あるいは学際的なテーマを設定し、個人研究に取り組み、4,000 語 (日本語の場合は 8,000 字) の論文にまとめるものである。CAS とは、以下 4 つの内容を含むさまざまな体験と少なくとも 1 つのプロジェクトに参加するものである。その 4 つとは、①重要な結果を伴う、真の目的をもった活動、②個人的な挑戦、③計画、進捗状況の確認、報告などの熟考、および④成果や個人的な学習についての省察である。

生徒が IB ディプロマを取得するためには、上記のカ

リキュラムを全て履修し、外部評価と内部評価を通じて、原則的に 45 点満点中 24 点以上を取得する必要がある⁹⁾。配点は、6 科目のそれぞれに 7 点 (計 42 点)、さらに TOK と EE の評価結果 (それぞれ、A～E までのいずれかの成績が出される) の組み合わせに応じて最大 3 点が加点される (表 1：なお、CAS は点数に含まれない)。

内部評価と外部評価の割合は科目によって異なる (表 2)。内部評価とは、例えば語学のオーラルワーク、地理のフィールドワーク、理科の実験、芸術的パフォーマンスなどを対象に、各 IB 認定校の教員によって実施されるものである。一方、外部評価とは、IB 機構が採点、評価する筆記試験や提出課題を指す。表 2 から、ほとんどの科目において、外部評価 70-80%/内部評価 20-30% という配分であること、ただしパフォーマンスの要素が大きい芸術科目では、内部評価の比重が大きくなることがわかる。換言すると、各科目の評価の基礎にあるのは、より高い客観性と信頼性が担保される外部評価であるといえる。なお、外部評価される試験において、多肢選択式の問題はまれであり、多くをエッセイ等の記述式問題が占める¹⁰⁾。

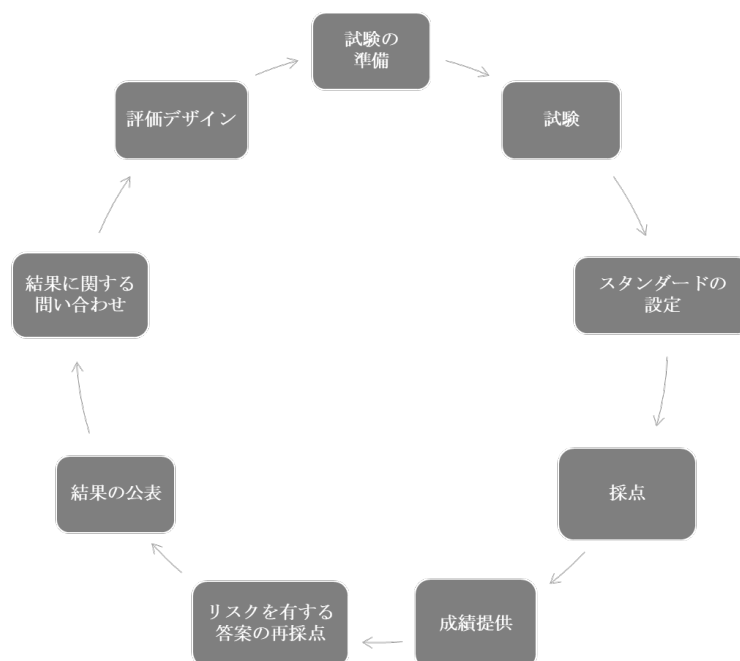


図4 IBの評価サイクル

出典：International Baccalaureate（2017a: 132）の図を筆者訳出

4. 外部評価の質の保証

4.1. IB における評価サイクルとその担い手

本節では、IB の外部評価の詳細をみていく。IB 機構は、評価の一連の手順を9段階から成る「評価サイクル」

（The assessment cycle：図4）と表現している。その9段階とは以下のとおりである。

- ① 試験の準備（Exam preparation）：個々の試験を作成するプロセス。出題テーマの決定から、具体的な問題の作成・編集、他言語への翻訳、正しい形式での配置・準備、そして最終的な質のチェックまで、全てを網羅する。
- ② 試験（Examinations）：受験者が自身の所属するIB認定校で評価を受ける。
- ③ スタンダードの設定（Standard setting）：上級採点官（senior examiners）が採点官チームに対して受験者の答案の採点方法を説明する。また、採点官の質をチェックするための確実な採点スクリプト（definitively marked scripts：詳細は後述）を特定する。
- ④ 採点（Marking）：採点官が受験者一人ひとりの答案を確認し、何点を付けるかを決定する。採点官は、主任採点官が出した指示に従わなければならない、それが正しく行われているかどうかをIB機構によって定期的にテストされる。
- ⑤ 成績提供（Grade award）：上級採点官が、点数（marks）

を成績（grades）に換算する方法を決定する。

- ⑥ リスクを有する答案の再採点（At risk re-marking）：採点の質の最終確認。問題がある可能性があることを示す証拠がある部分に焦点を当てる。
- ⑦ 結果の公表（Issue of results）：IB認定校と受験者に対する結果の公表。出された結果が受験者のプログラム合格（IBディプロマの取得）を意味するかどうかの判断や、大学への成績証明書の送付など、事務的な処理も含まれる。
- ⑧ 結果に関する問い合わせ（Enquiry upon results）：IB認定校に対して、試験の過程で誤りがあったと思われる箇所があった場合に、IB機構に受験者の答案をもう一度確認してもらえよう依頼する機会を提供する。
- ⑨ 評価デザイン（Assessment design）：評価サイクルを閉じるために最も重要な段階。受験者の経験から学び、評価課題の数や種類を含め、評価しようとする内容やアプローチを改善する。

この評価サイクルを1周回すことは、翌年の評価の質の向上につながるものであるとIB機構は考えている。

以下では、このサイクルのうち、特に公平性に関わると思われる③スタンダードの設定、④採点、および⑥リスクを有する答案の再採点について詳細にみていく。

4.2. 評価や採点に関わる主要なアクターの役割分担と責任の所在

IB 機構は、評価サイクルを担う主要なアクターとして、主任採点官、チーフ採点官、採点官、そして IB の 4 者を挙げ、それぞれの役割や責任を以下のように定めている (International Baccalaureate 2017a: 115)。

主任採点官

- ・ 評価のなかの一つの構成要素 (component) ⁽¹¹⁾ に責任をもつ。
- ・ その構成要素において、受験者の答案がどのような点数を取得するかを最終決定する。
- ・ その構成要素において、全ての採点官が採点基準 (marking standard) を理解していることを保証する。
- ・ その構成要素の成績換算基準 (grade boundaries : 点数 (marks) を 7 段階の成績 (grades) に割り当てる基準) を設定する際、チーフ採点官に対して指針を示す。

チーフ採点官

- ・ (主任採点官が特定の構成要素にフォーカスするのに対して) 担当する教科の全ての評価の構成要素を見渡す。
- ・ 試験問題とその問題の採点基準の確定 (ペーパーオーサリングと呼ばれる) を含めて、全ての評価の構成要素間で基準が一貫していることを保証する。
- ・ 評価に関連するあらゆる学術的問題を仲裁する。
- ・ IB に対して、最終的な成績換算基準を提案する。

採点官

- ・ 主任採点官によって設定された基準にしたがい、受験者の答案を採点する。

IB

- ・ 評価の全ての側面に責任をもつ。
- ・ 採点官のリクルートや採点官の質、試験結果の公表等、評価の一連のプロセスに責任をもつ。
- ・ 学業上の不正行為、不正管理、または特別な配慮についての決定を行う。
- ・ 上級採点官から提案された成績換算基準を受け入れる、またはそれに異議を申し立てる。

IB の採点の原則として、受験者の答案にどのような点数を付けるのが正しいかを最終的に判断する主体は主任採点官である、ということが挙げられる。主任採点官は、評価のなかの一つの構成要素に責任をもつことになるが、その担当 (どの構成要素の担当か) は一貫性を保つため

に複数回の試験に渡ることが通常である。なお、主任採点官は多くの場合、ペーパーオーサリングにも携わっている。

チーフ採点官は、IB とともに評価について議論し、改善する役割も果たす。主任採点官との違いとして、試験問題の質にも責任を負うことが挙げられる。なお、小規模な受験者数の科目において、IB はチーフ採点官を任命する代わりに試験責任者 (Examiner Responsible) を任命する。試験責任者は、科目全体の一貫性を確保するチーフ採点官の役割と、特定の構成要素に責任をもつ主任採点官の役割を兼任することとなる。

主任採点官やチーフ採点官は、IB のスタッフではなく、外部の専門家という位置づけである。外部の専門家との協働が、公平で質の高い評価を可能にすると考えられている。ただし、あくまで最終的な責任は IB 機構自体にある。

また、強い権限を与えられた、主に主任採点官の監督のもとで、個々の採点官は採点にあたることになる。

4.3. スタンドアードの設定と採点

上記の評価サイクルのうち、「受験者の努力に公平に報いることができるように、マークスキームをテストし、改良すること」を目的の一つとする段階が、スタンダードの設定である。スタンダードの設定は、以下 5 つを主要な目的としている (International Baccalaureate 2017a: 147)。

- ・ 受験者の答案に基づいて、主任採点官が (同僚の採点官との合議を通じて) 評価のためのスタンダードを設定すること
- ・ 受験者の努力に公平に報いることができるよう、マークスキームをテストし、改良すること
- ・ 確実な採点スクリプト (definitive mark scripts) を作成すること
- ・ 全ての採点官に理解を広め、共有すること
- ・ 採点官の「資格認定」(qualification) プロセスを通じて、スタンダードの理解を徹底すること

これらの目的の達成のため、スタンダードの設定は以下の手順で行われる。

はじめに、採点官のリクルートである⁽¹²⁾。適切な教育的背景やスキルを有すると思われる者が、採点官としてリクルートされる。そして、彼らが評価の手順や原則を理解できるよう、トレーニングが施される。

次に、標準化ミーティング (Standardization meeting) の

開催である。このミーティングでは、主任採点官をリーダーとする上級採点官 (senior examiner) のチームによって、マークスキームの作成がなされる。このマークスキームは、同一の回答に対して全ての採点官が同じ点数を付けることを可能にするために十分な情報を含むものでなければならないとされている。また、作成されたマークスキームが想定される範囲の答案を適切にカバーしていることを保証するため、試験終了後、上級採点官は受験者の答案例を確認することとなる。

さらに、標準化ミーティングの主要な成果の一つに、確実な採点スクリプト (definitively marked scripts) というものがある。これは、主任採点官によって採点済みの答案例であり、採点官をトレーニングしたりテストしたりするために用いられる。詳細な使用例は以下の3点になる。

第一に、典型的な答案をどのように採点するか、あるいはマークスキームを誤解しがちな状況としてどのようなものがあるか等を採点官が学ぶための使用である。採点官は、確実な採点スクリプトを読み込むことを通じて、正しく採点できるという自信をもつことが期待される。

第二に、採点官が正しい基準で採点していることを証明するための使用である。採点官は、IB 機構に対して正しく採点できることを示すことなしに、受験者の本番の答案を採点することはできない。正しく採点できるか否かのテストは、主任採点官によって採点済みの答案を採点することによってなされる。もし、主任採点官と「同程度の」点数を付けられなかった場合には、正しい点数となぜその点数を付けるのかについてのフィードバックが与えられる。そしてその後、採点基準の理解を示すための機会が改めて与えられる。なお、「同程度の」と表記する意図は、必ずしも主任採点官と完全に一致した点数を付けることを求められる訳ではなく、一定程度の幅が認められるからである。換言すると、同じ答案に対する正当な点数の差として「許容範囲」(tolerances) が認められている。

第三に、採点官が正しい基準で採点し続けていることを証明するための使用である。上記のとおり、採点を始める段階において、採点官は採点基準の理解や採点スキルをチェックされる。しかしながら、たとえ採点開始段階での理解等が保証されていたとしても、時間が経つにつれて採点基準にバラツキが生じる可能性は否定できない。そこで、各採点官が採点する答案のなかに、既に主任採点官が採点済みの答案を混ぜることによって(通常、その割合は 1/10 程度)、採点開始後にも採点の質の確認が行われる。結果として、付けられた点数と付けるべき

(主任採点官が付けた) 点数とが同程度であれば採点は続行されるが、もしそれらの間に乖離が見られる場合、IB 機構によるフィードバックとガイダンスがなされる。その上で、採点の再開を認められる場合もあれば、中止される場合もあるという。

このように、主任採点官と採点官とが同一の答案を採点し、その結果の差を比較することによって採点の質を保証しようとするアプローチは、品質モデル (quality model) とも名付けられている。このモデルが、記述式問題であっても、受験者がどの採点官に採点されたかにできるだけ左右されることなく公平な結果を受け取ることができるようにするための、つまり IB における外部評価の質保証のための核になっていると考えられる。

4.4. 採点ミスリスクを有する答案のレビュー

上記のとおり、IB 機構は、定期的に採点官の採点の質をチェックすることによって、外部評価の質を保証しようとしている。それに加えて、採点ミスリスクを有する答案のレビュー (“at risk” reviews) という段階が、採点の質の保証のために設定されている。以下、採点ミスリスクを有する答案のレビューとはどのようなものかを概説する (International Baccalaureate 2017: 191-192)。

採点ミスリスクを有する答案のレビューは、異常 (unusual) と思われる答案を探すことを目的とする。具体的な基準として、以下2つが設定されている。第一に、教員によって予想された成績⁽¹³⁾よりも実際の成績が著しく低い場合である。第二に、学校全体の成績が昨年と大きく異なる場合、とりわけ、評価のなかの特定の構成要素の成績が例年と大きく異なる場合である。これらに該当する場合、採点が必ずしも誤っているとは限らないものの、追加的な確認が必要であると考えられる。

採点ミスリスクを有すると判断された答案は、最も採点が一貫している採点官 (most consistent examiners) によって再採点される。なお、各採点官の採点の一貫性は、品質モデルの運用を通じて評価されている。再採点の結果が当初の結果と異なった場合、再採点後の点数が採用される。さらに、一貫性のある採点官 (consistent examiners) 2名の間においても大きな差が見られる場合、さらなる調査を実施し、なぜ差が生じるのかを理解するために第3者の意見を求めることもあるという。

以上のような手続きを設けることで、採点しにくい問題、あるいは採点に一貫性のない採点官等、外部評価の質を脅かす課題が発見される。そして、再採点を通じて、受験者ができるだけ公平な結果を受け取れるよう工夫がなされている。

5. 内部評価の質の保証

5.1. 内部評価の目的と課題

IBにおいて、内部評価は、学校（主に受験者を受けもつ教員）が受験者による課題への取り組みを評価し⁽¹⁴⁾、その後 IB 機構がその教員による評価の適切性（国際的なスタンダードを正しく適用しているか）を確認するという一連の手続きにより遂行される。なぜ、IB では、外部評価のみならず内部評価も重要視されるのか。内部評価の利点について、IB 機構は以下のように述べている。

「教室での評価（classroom-based assessment）は、試験にあまり適さない領域で受験者をテストする多くの機会を提供する。そのなかで最も重要なことは、受験者に対して、試験特有の時間的プレッシャーがない環境下で、問題を考察し、自分の思考をどのように発展させるかを示す機会を提供する拡張タスクに取り組むよう要求できることである」（International Baccalaureate 2017a: 94）。また、全世界統一の試験である外部評価だけでなく内部評価を設定することによって、IB 認定校は、地域、文化、または地理的な文脈のなかに学習を位置づけたり、教室と外の世界との結びつきを強めたりすることができるという。さらに、内部評価は、しばしば生徒に対して自身が関心をもつトピックや問題を選択する機会を提供するため、評価プロセスだけでなく学習経験全体の妥当性の向上に寄与するといわれている。

その一方で、内部評価については以下のとおりいくつかの問題点が指摘されている。第一に、受験者が不正行為（他者に作品をつくらせるなど）を行っていないことを確認することが非常に困難な点である⁽¹⁵⁾。第二に、内部評価のために設定された特定の課題に取り組むために、教員や生徒が過度に時間を費やし、結果として授業時間が圧迫されるリスクを孕む点である。

とはいえ、IB 機構は、受験者の成長を見てきた教員による教室での評価がもたらす価値は、それがもつ問題点を上回るという考えのもと、内部評価を実施してきた。

具体的には、教員は、以下3つの方法で IB ディプロマの授与に関わる評価に携わる。第一に、IB が示す基準（アセスメントクライテリア（assessment criteria）と呼ばれる⁽¹⁶⁾）に基づき、受験者が取り組んだ課題に対して内部評価の点数を付ける。第二に、IB が示す成績の基準（International Baccalaureate 2017b: グレードディスクリプター（Grade descriptors）と呼ばれる⁽¹⁷⁾）に基づき、各受験者が来たる最終試験で達成すると思われる成績を予想する。第三に、次項で説明するモデレーションと呼ばれ

る段階において、自身の採点基準の適用状況を確認されることとなるが、その際、受験者の課題に対してどのように内部評価したか（点数を付けたか）に関する説明責任を負う⁽¹⁸⁾。つまり、教員による評価といっても、IB 機構は彼らに完全に自由に評価させるのではなく、評価基準を課したりモデレーションを実施したりすることによって、質保証を図っている。

次項では、モデレーションの手続きについてより詳細に確認する。

5.2. モデレーション⁽¹⁹⁾

モデレーションとは、内部評価されたコースワークのサンプルに基づき、各 IB 認定校において実施された内部評価の点数を検証し、必要に応じて受験者の点数にモデレーション係数を適用するプロセスを指す。モデレーションにより、当該科目の内部評価を実施した全ての IB 認定校が出した点数を世界標準に合わせる可以考虑されている（International Baccalaureate 2023b: 91）。

より具体的には、外部採点官は、教員が評価したコースワークのサンプルを審査し、その教員が付与した点数が正しいか（すなわち、許容範囲（tolerance）内か）、厳しすぎる／甘すぎるか（すなわち、許容範囲外か）を判定する。教員が与えた点数が厳しすぎる、もしくは甘すぎる場合、モデレーション係数が決定され、その教員が属する学校の全ての点数に適用される。また、この際に学校は、自分たちの採点が（主任採点官によって設定される）世界標準とどう異なるかについてのフィードバックを受け取る。

また、モデレーションが公平であるためには、外部採点官が世界標準を理解し、それに基づきモデレーションに従事している必要がある。このため、モデレーションを担う採点官に対しても、外部評価の質保証のための品質モデルと同様の方法で（詳細は前節を参照）、その作業（モデレーションの適切性）に対するチェックがなされる。

モデレーション係数の適用は、上記のとおり学校を単位として行われる。このため、学校のなかで一つの科目に複数の教員が携わる場合、全員が採点を統一し、互いに矛盾がないようにする必要がある。採点の統一は、学校内で合意された一つの基準にしたがって全ての点数を付与することでなされる。さらに、最終的な評価を行う前に、教員間で話し合い、受験者が取り組んだ課題を合同でレビューすることも求められる。また、同一科目の標準レベルと上級レベルを異なる教員が担当する場合に

も、両レベル間で同一の合意された基準が適用されるよう、採点を標準化する必要があるとされる。

モデレーションに必要とされるサンプル数は以下のように定められている。すなわち、受験者が5名以下の場合、全員分/6~20名の場合、5名分/21~40名の場合、8名分/41名以上の場合、10名分である。なお、どの受験者の課題をモデレーションのためのサンプルとするかについては、学業上の不正行為を排除するために、IB機構によって指定される。

IB機構は、モデレーションにおけるサンプリングにあたり、以下2つの原則を定めている。第一の原則として、IB機構は、特定の学校の信頼できるモデレーション係数を得るために必要最小限のサンプルを使用する。これは、学校の負担を最小限にするため、また、費用を抑えるためである。ただし、信頼できるモデレーション係数を決定するために、より多くの教員の採点例を必要とする場合、IB機構は各IB認定校に対して追加のサンプル提出を依頼することがある。このため、各学校は、結果の公表段階まで全ての受験者の課題を入手できるようにしておく必要があるとされる。第二の原則は、モデレーション係数（の掛け方）があらゆる得点範囲にまたがる全ての受験者にとって公平でなければならない、ということである。それゆえ、モデレーションのサンプルは、全ての得点範囲（高得点群、中得点群、低得点群）が適切に代表されるよう選択されなければならないとされる。というのも、教員は、採点範囲内の異なるポイントで異なる期待をもつことがあり、質の低い課題に対して世界標準より寛大な教員が、質の高い課題に対して世界標準より厳しい可能性があるためである。さらに、満点を獲得した受験者はモデレーションサンプルに選ばれない傾向がある。これは、教員の採点が厳しすぎる場合、高得点群の受験者を適切に上方修正するためである。

以上、本節では内部評価の質の保証がいかに図られているかを整理した。質の保証の核には学校を単位とするモデレーションが位置づけられるものの、果たしてそれは本当に公平な評価につながるのか。この点については次節第3項で触れることとする。

6. IBの評価における公平性をめぐる課題

以上のとおり、IB機構は評価の公平性の確保に努めている。とはいえ、課題を認識していないわけではない。以下では、(1)試験が国際的なものであるがゆえの課題、および(2)生成AIが台頭している現代に特有の課題に焦点をあてて、IB機構の課題認識と対処、議論の

現在地を示す。加えて、評価システムの根幹に不確実な（外部からの検証が難しい）仮定を置いていることの問題を指摘する。

6.1. 試験が国際的なものであるがゆえの課題

IBによる試験と国を単位として実施される大規模共通試験の決定的な違いは、受験者の居住地や文化的背景の多様性の程度にある。本項は、その多様性の大きさがIBの評価の公平性に対していかなる課題を突きつけているかを示す。

第一に、受験者の居住地が国際的に分布しており、ゆえにタイムゾーンをまたいだ試験実施が不可避であることに由来する課題である。国を単位として試験が実施される場合、基本的に同一時刻に試験が開始され、終了する。一方、IBのように複数国で試験が行われる場合、時差のために開始・終了時刻を統一することは事実上不可能である。ゆえに、試験問題の漏洩リスクは高まらざるを得ない。この問題について、IB機構は、「国際的な試験や時差の問題への挑戦」という項目をガイドラインのなかに設け、その対策について以下のように論じている（International Baccalaureate 2017a: 123-124）。

対策の一つ目に挙げられるのは、「利己主義」（self-interest）という考え方、つまり、受験者は自分よりも他の受験者に利することを選択しないだろうと想定することである。二つ目は、問題用紙が共有される可能性のあるウェブサイトの注意深い監視である。また、問題用紙を試験会場から持ち出すことは許されないという規則が、インターネット上で問題が共有されるリスクを制限するために設けられている。三つ目に、知識の想起ではなく理解を問うような問題を作成することが挙げられている。このことは、思考のスピードを評価対象としないような出題をすれば、事前に（試験開始の数時間前に）問題を知ることによって回答を考える時間が少々延びることから得られる利益は限定的であろうという想定に基づいている。四つ目が、全ての受験者に対して試験終了後24時間は試験について話してはならないという規則を課す、というものである。IB機構は、以上の方策を示したうえで、この課題に対する議論を今後も継続することを明言している。

さらに、言語的・文化的な中立性の確保も、公平性の問題として指摘できる。Badham (2025) は、「国際的評価が抱える主な課題」と題した記事のなかで、国際的多様性をもつIBの試験は固有の難題を有することを指摘し、「異なる言語間で難度を等しく保つためにはどうしたらよいか」「文化的に包摂的な試験問題をどのようにつくる

か」「評価基準（例：適切な（adequate）、よい（good）の翻訳差をどのように調整するか」といった問いを提起している。同様の課題意識を、IB 機構も共有しており、ある文化圏において馴染み深い文脈に沿った問題が他の文化圏の者にとっては異質で理解しにくい場合、その試験問題は偏り（bias）をもっていること、「分析せよ（analyse）」「評価せよ（evaluate）」「説明せよ（explain）」といったコマンドターム（command terms）が翻訳される際、その用語がもつ意味の微妙なニュアンスが消失してしまう可能性があること等を課題として指摘している。そしてそれらの課題への対策として、問題文のなかで、問題の理解に必要な背景情報を提供することや生徒が自身の文化的文脈を応用して解答できるような、よりオープンエンドな問題文を設定すること、各言語におけるコマンド用語とその意味を公表・解説することを、IB 機構は行っている（International Baccalaureate: 2017a）。さらに、より中立な試験、評価の実施に向けた調査研究が進展していることも報告されている（Badham 2025）。言語・文化を超えた中立性の確保は、IB 機構が自覚し取り組んでいる、評価をめぐる中核的課題の一つとしてとらえられるだろう。

6.2. 生成 AI が台頭する現代に特有の課題

さらに本項は、現代に特有の課題として、生成 AI の台頭をいかに IB 機構がとらえているのかに触れる。

IB 機構は、近年、「評価と教育における ChatGPT ならびに人工知能（AI）の役割」⁽²⁰⁾と題する声明を出し、加えて「IB の評価と教育における人工知能（AI）：危機か機会か」⁽²¹⁾と題する記事のなかで、生成 AI への立場性を示してきた。また、Academic integrity policy（International Baccalaureate 2023d）のなかで、生成 AI の利用に関するガイドラインを示してきた。それらの特徴として、①生成 AI への対応は、従来の不正行為（外部チューターや家族による代筆）の延長線上に位置づけられ、ゆえにゼロから対策を考えるのではなくこれまでのポリシーの発展を要求するものと考えていること、また②生成 AI は電卓や翻訳ソフトなどと同じく日常化するツールであり、かつ脅威でなく知的探究を深化させるものだという将来像を描いていること、ゆえに③AI の問題に対して、利用の制限・禁止でなく、生徒への教育的介入を通じて対処しようとしていることを挙げられる。以下、International Baccalaureate（2023d）に依拠して、生成 AI の課題に対する IB のアプローチを詳述する。

大原則として、AI が生成したいかなる成果物も、生徒（受験者）自身のオリジナルの著作物とはみなされない。

ゆえに、生成 AI を利用した際には、学術論文などの引用と同様に、出典を示すことが求められる。具体的には、AI が生成したテキストを直接引用したり、パラフレーズして引用したりした場合、当該箇所が生成 AI によるものであることを本文中で明確に示し、かつ利用した AI ツールの名称を参考文献リストに記載することが求められる。これらの責任を果たさなかった場合、学問的不正行為として扱われる。生成 AI の利用に対しても、知的成果の帰属を明確にするという学問の基本原則が適用されるという方針が示されている。

そもそも、IB の評価において生成 AI の利用が問題になる場面は、主に各 IB 認定校における内部評価である。ゆえに、特定の課題が生徒自身によるものか否かを判断する責任は、原則的に各教員にあると考えられる。教員には、学習成果の真正性を担保する最終的な責任が課せられているのである。真正性を判断するための具体的観点として、①長期的な制作過程の観察、②口頭説明の整合性、③適切な引用、④成果物の水準と生徒自身の学力の一致という 4 要件が示されており、本人の取り組みであることについて疑義が解消されない場合、教員は当該取り組み（作品）を IB 機構に提出しない義務を負う。

以上で示したような、生成 AI の禁止でなくそれとの共存を志向するアプローチは、先進的であると評価できるかもしれない。しかしながら、内部評価への取り組みの過程における生成 AI の活用に関する指導も、学習成果の真正性の判断も、各 IB 認定校の教員の裁量に委ねられている点で、不確実性を伴う。とりわけ、学習成果の真正性の判断、つまり、ある課題が本当に生徒自身によるものか（AI を活用したものでないか）を教員が判断することは容易でない（Fleckenstein, Meyer, Jansen, Keller, Köller & Möller 2024）。例えば、学業的に急な成長をみせたり、あるいは瞬間的にアイディアを閃いたりするタイプの生徒がいたとして、彼らが誤って「生成 AI を用いた」と判断されてしまうリスクはある。このリスクを踏まえると、生成 AI の利用の有無の判断を現場（教員）に任せることは、酷であるとも評せる。生成 AI の利用をめぐる指導や判断を現場に委ねることの不確実性が、IB の評価の公平性の揺らぎにいかんに影響を及ぼすのか否かについては、今後注視すべき点の一つである。

6.3. 評価システムの仮定が内包する問題

IB の評価の公平性を確保する仕組みは、その土台に、外的に検証困難な仮定を置いている。このことが、公平性をめぐる構造的な問題を内包していることを、本項では指摘したい。

その仮定とは、第一に、「主任採点官の採点は正しい」というものである。第4節、第5節で整理したとおり、IBにおける公平性の確保の核には、主任採点官の存在がある。換言すると、「主任採点官の採点は正しい」という仮定のもと、品質モデルの運用やモデレーションの適切性の確認がなされている。採点官の出す得点が必ずしもそのまま採用されるとは限らず、採点の正しさの確認が定期的になされている点は、記述式問題の採点を公平に行うための優れた取り組みのように思われる。しかし、上記の仮定の根拠となる、主任採点官の選任基準に関しては、2015年発行の資料から概要を確認できるものの（International Baccalaureate 2015）、最新の状況を把握することは難しい。加えて、主任採点官に対する具体的なトレーニングの有無や内容、および主任採点官の評価の確からしさ（例えば、主任採点官同士の評価者間信頼性）などについては、明らかにされていない。言い換えると、彼らの採点は正しいということはいかなる根拠に基づいているのかという点について確認することが困難な現状がある。このことは、評価システム全体の透明性、公平性に疑念を抱かせる要因になりうる。

第二の仮定は、モデレーションに関わるものである。IB機構は、学校を単位としてモデレーションを実施しており、それにより学校間格差（学校による採点の甘さ／厳しさの差）の問題にうまく対応しているようにも思える。学校を単位とするモデレーションの実施の前提には、「学校のなかで一つの科目に複数の教員が携わる場合、全員が採点を統一している」という仮定がある。果たして、この仮定は全ての学校で成り立つのか。なお、モデレーションが評価結果（評点）の平準化の過程を指すのに対して、評価者間の判断（物差し）を統一化する過程はキャリブレーションと呼ばれる（石田 2021）。つまり、IB機構による内部評価の質の保証の核にあるモデレーションを適切に実施するためには、前提として適切にキャリブレーションが実施されていなければならない。ところが、適切にキャリブレーションが行われているかという点についても、外的には検証が難しい。IB機構は、学校内で合意された一つの基準にしたがって全ての教員がコースワークの評価にあたるという指針を示しているものの（International Baccalaureate 2017a）、その実効性は学校の状況、文脈に依存する可能性が高い。結果として、「モデレーションのサンプルに含まれなかった生徒が、正確な（得べき）得点を得ていたにもかかわらず、他者の採点結果を根拠として不当に得点を引き下げられる」といった避けるべき事態が実際に起きているか否かについて、判断し難い。それゆえに、IBの評価の公平性に疑

念を抱く余地が生じることとなる。

このように、「主任採点官の絶対的な正しさ」と「各学校における適切なキャリブレーション」という二つの仮定（の不確かさ）は、IBの評価システムが意図せざる不公平を生み出す危険性を孕んでいることを示唆している。

7. まとめ

IBにおいては、「優れた評価は、生徒の幅広いコンピテンシーと高次の思考力を考慮したものである」、「知識、概念、ルーチンのテクニックを思い出すことだけに頼いるテストは、IB教育の目標にそぐわない」（International Baccalaureate 2017a: 97-98）という考え方のもと、記述式問題を中心にした外部評価、および各IB認定校の教員による内部評価が実施されてきた。本稿は、それら外部評価と内部評価において、いかに評価の公平性の確保が試みられているのかを明らかにすることを目的として、IB機構が公表しているガイドラインを読み解いてきた。加えて、評価の公平性に関わるいくつかの課題を指摘してきた。得られた知見の要点は以下のとおりである。

外部評価の質保証は、単に採点基準（マークスキーム）を設定するのみならず、品質モデルの運用によって試みられている。すなわち、主任採点官による採点結果は正しいという仮定のもと、主任採点官と（その他の）採点官の採点結果を比較することにより、採点官としての資格を与えたり、採点の過程でその資格を維持させるか否かが決定されたりする。さらに、採点が難しい答案をリスク（at risk）にとらえ、それについては特に採点が一貫している採点官によって慎重に検討される段階が設けられている。

内部評価の質保証についても、同様に、IB機構の教員に対して評価基準（アセスメントクライテリア）を課すだけでなく、モデレーションと呼ばれる段階が設定されている。つまり、各IB認定校で行われた内部評価の適切性が、ランダムに選ばれた評価サンプルによって確認され、場合によっては成績の調整がなされる。

以上の多層的な手続きによって、記述式問題あるいはコースワークの評価であっても、一定程度、公平な評価が実現されていると解せる。IBに対しては、その設立当初より、そもそも評価の客観性や公平さを保証するなにかが存在するかが存立そのものに関わる問題であるということが指摘されてきた（西村 1989）。このような指摘に対してのIB機構による今日的な応答の一端として、本稿が整理してきた評価の仕組みを位置づけることができる。

他方で、IB 機構も自覚しているとおおり、試験が国際的に実施されるがゆえの課題や生成 AI の活用をめぐる課題は、即時的にあるいは容易に解消されるものではない。ゆえに、IB の評価を完全なもの、あるいは完成されたものとみることは誤りである。IB というシステムに参画する学校、国、地域の増加・多様化、あるいは生成 AI の台頭に顕著に示される社会変化に、IB 機構はいかに対応し、今後、評価システムを発展させていくのか。加えて、「主任採点官の絶対的な正しさ」や「各学校における適切なキャリブレーション」という仮定の確かさに関して、いかに対外的に証明していくのか。それらの点を注視していくことが、IB の評価の公平性を考察するためには必要不可欠であり、研究上の今後の課題としても位置づけられる。

<注 記>

本稿は、科学研究費補助金基盤研究(B)「大学入学者選抜統一試験における制度変容と公正知覚に関する国際比較研究」(研究課題番号: 22H00978/23K22249、研究代表者: 九州大学教授 木村拓也)による研究成果の一部である。

<注>

(1) 国際バカロレア (IB) は、ディプロマプログラムに加えて、初等教育プログラム (Primary Years Programme)、中等教育プログラム (Middle Years Programme)、およびキャリア関連プログラム (Career-related Programme) で構成されている。ただし、本稿では、記述の煩雑さを避けるため、「IB」を IB ディプロマプログラムの略称として使用する。

(2) IB 機構(ジュネーブに拠点がある IB 本部)によって、IB プログラムのうちのいずれかの実施を認可された学校を指す。

(3) IB 機構が、いかに IB ディプロマの認知を獲得してきたかについては、Resnik (2016) に詳しい。

(4) IB 機構は、公平な評価の目的について、全ての受験者に等しい機会を提供することであると述べている (International Baccalaureate 2017a: 68)。

(5) 本稿が検討している IB 関連の文書は、全てインターネット上で一般公開されているものである。IB 機構は、IB 関連の文書の多くを My IB と呼ばれるウェブサイト (<https://ibo.my.site.com/ibportal/IBPortalLogin?startURL=%2Fibportal%2F>) において共有しているが、アクセス権限は IB 認定校の教員など一部の者に限定されている。ゆえに、そこに掲載されている文書は本稿の検討の対象か

ら外れる。

(6) □ 内は引用者による。

(7) コーディネーターとは、IB プログラムの実施においてリーダー的な役割を果たし、加えて自身の所属する IB 認定校と IB 機構との連絡調整を担当する教員を指す (International Baccalaureate 2023c)。

(8) International Baccalaureate, <https://www.ibo.org/> (accessed 2025-6-24)。

(9) ただし、いずれかの科目で 1 点を取った場合等には全体で 24 点を超えていても IB ディプロマは授与されないなど、細かな資格授与の規則が定められている (International Baccalaureate 2017a: 220)。

(10) 例えば、歴史では「産業化が 1 つの国の生活水準と労働環境に及ぼした影響について考察しなさい」といった出題がなされる (International Baccalaureate 2021a)。

(11) IB の各科目の評価は、例えば外部評価の対象となる筆記試験 1 (paper one)、筆記試験 2 (paper two)、口述試験、および内部評価から成る。それらが評価の構成要素と呼ばれる。

(12) IB 機構はリクルートのためのポリシーを公表しており (International Baccalaureate 2021b)、そこで定められている条件を満たしかつ採点官になることを希望する者が応募できる仕組みになっている。

(13) IB では、最終試験に先立ち、IB 認定校の教員によって予想される成績 (predicted grades) が出され、大学への出願等に際して活用される。

(14) 内部評価の手順は、ガイドライン (International Baccalaureate 2023b) 中の「内部評価と予想された成績」 (Internal assessment and predicted grades) の項目で詳しく説明されている。この項目は、点数の妥当性と信頼性を保証するために設計されており、内部評価に関わる教員が熟知しておくべきとされている。

(15) 不正行為への対応については、IB 機構によって、Academic integrity policy というガイドラインが出されており (International Baccalaureate 2023d)、そのなかで不正行為一覧 (盗用、共謀、外部委託・答案購入、重複提出、データ捏造・改ざん、試験における不許可物所持、試験における情報のやりとり、試験問題の事前入手・漏洩、代替受験、不正行為の不申告など) とそれらに対する罰則が列挙されている。

(16) アセスメントクライテリアについては、各教科のガイドで詳細が確認できる。各教科のガイドは、IB 認定校の教員等によってアクセス可能な My IB 内のプログラムリソースセンターで入手可能であるが、一般公開はされていない。ただし、特定のウェブサイトでのみ

(<https://www.follettibstore.com/>) 購入可能である。

(17) グレードディスクリプターについては、IB 機構によって 2017 年に出された文書 (Grade descriptors For use from December 2017) で詳細が公開されている (International Baccalaureate 2017b)。

(18) モデレーションの過程で、教員は、なぜそのような評価をしたのかについて、詳細な説明を求められることがある (International Baccalaureate 2017a: 168-169)。ただし、

(19) 本項の記述は、主にガイドライン (International Baccalaureate 2023b: 91-94) に基づいている。

(20) International Baccalaureate, “Statement from the IB about ChatGPT and artificial intelligence in assessment and education.” <https://www.ibo.org/news/news-about-the-ib/statement-from-the-ib-about-chatgpt-and-artificial-intelligence-in-assessment-and-education/> (accessed 2025-6-30).

(21) International Baccalaureate, “Artificial intelligence in IB assessment and education: a crisis or an opportunity?” <https://www.ibo.org/news/news-about-the-ib/artificial-intelligence-in-ib-assessment-and-education-a-crisis-or-an-opportunity2/> (accessed 2025-6-30).

<参考文献>

有馬実世, 2022, 「国際バカロレア中等教育プログラムの単元設計に取り組む教師の評価観の変容のプロセス—形成的評価に注目して—」『国際バカロレア教育研究』6, pp.95-104.

Badham, Louise, 2025, “Challenges in international assessment.” Retrieved from <https://www.ibo.org/research/assessment-research/research-studies/challenges-in-international-assessment/> (accessed 2025-6-30).

Bunnell, Tristan, 2011, “The growth of the International Baccalaureate Diploma Program: Concerns about the consistency and reliability of the assessments.” *The Educational Forum*, 75(2), Taylor & Francis Group, pp.174-187.

Fleckenstein, Johanna, Meyer, Jennifer, Jansen, Thorben, Keller, Stefan, D, Keller, Olaf, Köller, & Möller, Jens, 2024, “Do teachers spot AI? Evaluating the detectability of AI-generated texts among student essays.” *Computers and Education: Artificial Intelligence*, 6, 100209.

Hill, Ian, & Susan, Saxton, 2014, “The International Baccalaureate (IB) programme: An international gateway to higher education and beyond.” *Higher Learning Research Communications*, 4(3), pp.42-52.

International Baccalaureate, 2015, Principal examiner recruitment policy.

International Baccalaureate, 2017a, Assessment principles and practices—Quality assessments in a digital age.

International Baccalaureate, 2017b, Grade descriptors For use from December 2017.

International Baccalaureate, 2020a, Assessment principles and practices—A guide to assessment for students and parents/guardians.

International Baccalaureate, 2020b, Assessment principles and practices—A guide to assessment for teachers and coordinators.

International Baccalaureate, 2020c, Assessment principles and practices—A guide to assessment for examiners.

International Baccalaureate, 2020d, Assessment principles and practices—A guide to assessment for chief examiners and principal examiners.

International Baccalaureate, 2021a, 「国際バカロレア (IB) ディプロマプログラム (DP) 科目概要「個人と社会」: 「歴史」—標準レベル (SL) 2017 年第 1 回試験—2025 年最終試験」.

International Baccalaureate, 2021b, IB Examiner recruitment policy Version 1.6 01/07/21.

International Baccalaureate, 2022a, The IB Diploma Programme and Career-Related Programme Final Statistical Bulletin May 2022 Assessment Session.

International Baccalaureate, 2023a, The IB Diploma Programme Final Statistical Bulletin November 2022 Assessment Session.

International Baccalaureate, 2023b, Diploma Programme Assessment procedures 2023.

International Baccalaureate, 2023c, Considerations for selecting your programme coordinator.

International Baccalaureate, 2023d, Academic integrity policy.

International Baccalaureate, 2024, The IB Diploma Programme and Career-Related Programme Final Statistical Bulletin May 2024 Assessment Session.

International Baccalaureate, 2025, The IB Diploma Programme Final Statistical Bulletin November 2024 Assessment Session.

石田智敬, 2021, 「モデレーション／キャリブレーション」西岡加名恵・石井英真編『教育評価重要用語事典』明治図書出版株式会社、p.52.

中村泰輔・小笹哲夫, 2022, 「国際バカロレア (IB) 日本語デュアルランゲージ・ディプロマプログラムにおける理科学科の実践」『理科教育学研究』63(1)、pp.3-13.

西村俊一, 1989, 「世界の国際学校と国際バカロレア」西村俊一編『国際的学力の探究 国際バカロレアの理念と課題』創友社、pp.33-48.

Resnik, Julia, 2016, “Struggling for recognition: access to higher education through the International Baccalaureate.” *Critical Studies in Education*, 60(3), pp.340-357.

Tan, Sierra, n.d., “Growing Pains: Challenges and Opportunities for the International Baccalaureate.” Retrieved from <https://bpb-us-e2.wpmucdn.com/sites.middlebury.edu/dist/8/6018/files/2013/06/Growing-Pains-of-the-IB.pdf> (accessed 2025-6-23).

<付記>

本稿は、以下の既発表の報告書に加筆・修正を加えたものである。

江幡知佳, 2023, 「国際バカロレアの評価における公平性の確保」木村拓也編『海外の大規模大学入学共通試験における公平性ポリシー・ガイドラインに関する調査研究』独立行政法人大学入試センター研究開発部、pp.113-122.