

個別入試における作題関連業務を規定する要因の探索的分析と 仮説モデルの構築¹⁾

—作題支援システムの構築を目指して—

久保 沙織 (東北大学)

本研究の目的は、良質な問題の効率的かつ継続的作成に資する作題支援システムの構築を目指すために、個別試験の作題関連業務に従事する大学教員への面接調査のデータから、業務を規定する要因を探索的に抽出し、その背後に仮定される認知プロセスに関する仮説を生成することである。作題関連業務はその手順によって、①作問、②配点・採点基準の設定と採点、③試験実施後の評価、の3つのフェーズに分類でき、各フェーズを規定すると考えられる複数のカテゴリが抽出された。認知プロセスの仮説モデルにおいては、学問分野による違いに加えて、個人のパーソナリティや経験に依拠する信念も一定の影響を持つ要因であることが示唆され、作題支援の具体的方策についても明確になった。

キーワード：個別入試、作題、作題支援、認知モデル、仮説生成

1 問題・目的

大学入試の唯一かつ重大な目標は、アドミッション・ポリシーに表現された「求める学生像」に沿った学生を定員通りに確保することであり(倉元, 2020)、入試問題の作成は、この目標を達成するために個別大学に課された責務と言えよう。さらに、大学入試は受験生の将来や意思決定に大きな影響を及ぼすことから、日本社会におけるハイスタークスの試験の典型であり、入学者選抜のための試験問題は、その質が問われる。一方で、作題関連業務の過重負担、短期間での担当者の入れ替わり、入試ミスといった様々な阻害要因があり、良質な問題の継続的作成は、多くの大学にとって深刻な課題となっている。

大学入試における個別学力検査の質の評価を目的とした実践例を報告している先行研究の1つとして、木村(2012)が挙げられる。木村(2012)では、合格者を対象としたアンケート調査によって得られた受験者の主観的評価と、実際の学力検査の得点との関係性に基づく作題評価の可能性について論じている。

しかしながら、入試問題の作成は高度な秘匿性が求められる業務であり、言わば聖域として扱われているため、これまでその業務内容が詳に語られることはなかった。このような事情もあり、作題関連業務の一連のプロセスを明らかにした上で課題を特定し、具体的な支援方法について議論している研究はほとんど見当たらない。佐藤ほか(2013)では、作題のPDCAサイクル(plan-do-check-act cycle)の構築に向けて、作題者へフィードバックする内容を検討した新潟大学の取り組みが報告されているが、当該大学の入学センター

でピックアップした内容の分析例の提示にとどまっておき、そのフィードバック内容が実際に有用であったかに関する作題者からの評価までは検討されていない。

本研究では、作題関連業務に携わる大学教員への面接調査を通して業務の実態を把握し、当該業務を規定する要因を探索するとともに、その背後に仮定される認知プロセスを明らかにすることを目的とする。なお、本研究の目的は、各大学が一般選抜の問題に関して公表している作題意図のように、出題された個々の問題に関する振り返りや個別大学が入試問題に込めたメッセージを紐解くことではない点に留意されたい。本稿の成果を礎に、大学入試における個別試験の作題負担を軽減し、かつ、ハイスタークスの試験にふさわしい良質な問題を効率的かつ継続的に作り続けるための作題支援システムの構築を最終目標とする。

作題支援の方法としては、大きく2つの方向性が考えられる。1つは、FD等を通して作題関連業務に必要な知識や技術を教授することで、作題者の能力向上を図る情動的側面からの支援である。もう1つは、作題のために有用なデータベース等を作成し、その活用により負担の軽減や作業の効率化を図る道具的側面からの支援である。本研究では、これらを含めて作題支援システムと称している。作題関連業務における認知プロセスを一般化した仮説モデルによって表現することで、課題の所在と必要とされる支援が明確化され、実用的な作題支援システム構築の第一歩となることが期待される。

2 方法

2.1 調査対象者および調査時期

調査対象者は、某大学で作題関連業務に携わる経験を持った教員 6 名であった（教員 A, B, C, D, E, F とする）。各教員の当該大学における作題関連業務への従事年数、および、専門分野を表 1 に示した。なお、6 名とも男性であった。

調査時期は 2021 年 3 月中旬から下旬にかけてであった。

表 1 調査対象者の従事年数と専門分野

教員	従事年数	専門分野
A	5 年	理科 (物理)
B	4 年	理科 (化学)
C	2 年	理科 (生物)
D	5 年	数学
E	4 年	英語
F	1 年	国語

2.2 半構造化面接

上記対象者に対して 1 人ずつ、半構造化面接 (semi-structured interview) によるインタビュー調査を実施した。

2.2.1 質問項目

質問項目は、本学入試センターでの業務経験に基づいて著者が考案し、センター内の他の教員と合議の上、決定した。当初 12 の質問項目を用意し、実際に質問を行ったが、本稿の目的に鑑み、分析対象とした項目は表 2 に示した 7 つとした。

表 2 質問項目

項目 1	試験問題を作る際、素材集め、テーマ決めなど、どのようにしているか？
項目 2	採点基準はどのように決めているか？難易度の予想や調整はどのようにしているか？
項目 3	作った問題が意図した通りに機能したなどと思うとき、そうでないときでは何が違うと感じているか？
項目 4	出題した問題を検証したり、その結果を次年度以降の作題に活かすために、どのような視点で分析を行っているか？
項目 5	入試問題を作る上で、大切にしていることはどんなことか？
項目 6	作題支援関連業務で、一番苦勞する点、辛いこと、困っていることは何か？
項目 7	個別大学の入試問題における良い問題とは？悪い問題とは？

2.2.2 手続き

調査対象者には、面接を実施する前に研究の目的および概要、プライバシー保護や研究データの取り扱い等に関して説明し、同意を得た。面接時間は 60~150 分であった。面接の内容は、調査対象者の了承を得た上で、IC レコーダーで録音した。面接終了後に、録音された内容を書き起こし、逐語録を作成した。

3 結果と考察

3.1 カテゴリ抽出

逐語録から、一義的に解釈できる意味内容ごとに断片化した結果、255 のデータが得られた²⁾。著者と、心理学を専門とする研究者である第三者の 2 名での合議の下、カテゴリ分類を行った。まず、個別試験の作題関連業務は、その手順によって、①作問、②配点・採点基準の設定と採点、③試験実施後の評価、の 3 つのフェーズに分類できることで合意を得た。これら 3 つを大カテゴリとし、それぞれの作業フェーズにおけるキーワードをもとに小カテゴリを作成し、それらを概念的にまとめたものを中カテゴリとした。

各フェーズで抽出された中カテゴリおよび小カテゴリ、そして対応する回答例を表 3 から表 5 に示した。回答例の括弧内には、調査対象者を識別する記号と専門を記し、それぞれの回答がどの対象者のものかを明示した。なお、表中の回答例については、個人情報保護および作題業務の秘匿性の観点から、文意を損なわない範囲で筆者が加筆修正を行っている。中でも、「配点・採点基準の設定と採点」は、特に秘匿性の高い業務であると判断し、表 4 には回答例は記載せず、それぞれの小カテゴリに対して言及のあった対象者を示すにとどめた。表 3 から表 5 の記載内容について、発話意図と相違がないか調査対象者の確認を得た。

3.2 カテゴリに基づく考察

以下では、中カテゴリは【 】, 小カテゴリは〈 〉, 発話内容は『 』で表す。

3.2.1 「作問」のフェーズ

作問のフェーズにおいてまず行われるのは、理科 (物理, 化学, 生物) ではテーマ・素材の決定、英語と国語では素材文の選定、数学では問題そのものの選定であることが示された。その際の拠り所や参照する対象を表すカテゴリである【素材の収集源】では、〈教科書〉〈本〉〈他大学の問題〉〈新聞・雑誌〉というカテゴリが抽出された。

「作問」のフェーズでは、教員ごと、あるいはその

教員が専門とする教科・科目ごとの違いが特に明確に見られたため、以下では、教員ごとに発言内容を整理して考察する。

教員 A・物理 物理では、【問題のタイプ・分類の認識】にあるように、問題で扱うテーマの異同は、『素材として使われる図』によって弁別される点が特徴的であり、化学や生物と同様にテーマから決定する方略であると見なすこともできるが、素材文の選定が最初に行われる英語や国語との類似性も見出された。教員 A は【重視する観点】として、目標とする平均点を意識した〈難易度〉と、可否に関する〈識別力〉を挙げており、良い問題に関する発言内容から、【測定したい資質・能力】は〈思考力〉であることが推察された。

教員 B・化学 【重視する観点】のうち、教員 B の発言にあった〈らしさ〉は、AO 入試や一般選抜といった選抜の種類ごとの目的、あるいは学部の特徴に合致していることを指し、内容的妥当性に通ずる概念である。それ故に、作題時の『回答形式の変更による得点調整』や、実施後の『平均点調整』等による恣意的な難易度の調整は、〈妥当性〉を損ねるものとして忌避されている。教員 B は、作題において最も大切にしていることは〈メッセージ性〉であると述べており、その意味するところは大学として『何を学んできてほしいかを明確に打ち出すこと』、すなわち【測定したい資質・能力】として〈大学が求める資質・能力〉が何であるかが明確に伝わる問題であることを重要視していると解釈できる。その具体例として〈思考力〉〈応用力〉が挙げられており、それらを測定できる問題は良い問題であると考えられている。また、〈記述式問題〉がそれらの資質・能力を有効に測定するための問題形式の 1 つとして捉えられていることが、【問題のタイプ・分類の認識】で示唆されている。

教員 C・生物 教員 C は、【重視する観点】として、生物というディシプリン、すなわち〈学問観〉を大切にしていることが推察された。また、【問題のタイプ・分類の認識】では、生物における〈テーマ〉とは、『素材ではなく、実際に起こっていることのカテゴリ』であり、それらのカテゴリは『非常に限定的であり、かつ互いに関係していて明確に切り分けようがない』と述べられていた。そのため、【重視する観点】として他大学や共通テストと被らないという〈重複回避〉が重要になると考えられる。また、【測定したい資質・能力】に〈思考力〉や〈幅広い視野〉が挙げられていることの根底にも、上述のような生物という学問分野の特質が関連しているのではないだろうか。

教員 D・数学 教員 D は、〈教科書〉をベースに、問題集や〈他大学の問題〉等を幅広く参照し、【重視する観点】にあるように〈教科書・問題集との類似性〉を意識しつつ、教科書を基準として〈難易度〉の判断を行っている。さらに、『良い問題とは、高校の教材として使ってもらえるような問題』と回答していることから、〈教材としての価値〉によって問題の質を評価していると解釈することができる。また、作問において〈難易度〉と〈識別力〉を意識しており、特に〈難易度〉には【問題のタイプ・分類の認識】における〈小問形式・大問形式〉の違いが関係していると捉えられていた。

教員 E・英語 英語を専門とする教員 E の発言内容からは、素材文ありきの認知プロセスが窺える。【重視する観点】から、素材文ごとに〈素材文の語数〉と〈難易度〉を記録し、選定のための材料としていること、素材文の内容に込められた〈素材文のテーマの面白さ〉や〈メッセージ性〉を重視していることが示された。また、〈採点のしやすさ〉は、信頼性の確保を示唆しているものと考えられる。〈入試ミス〉を出さないという実務的な観点は、他の教員には見られない特異な視点であった。【問題のタイプ・分類の認識】では、〈語順並べ替え問題〉〈説明問題〉〈英文和訳問題〉といった問題形式と、測定し得る能力がセットになって認知されていることが明らかとなった。

教員 F・国語 教員 F は『素材文を選んだら半分以上決まると言ってもいい』と述べており、英語と同様に、国語においても素材文の選定が第一であり、〈素材文のテーマの面白さ〉や素材文の内容の持つ〈メッセージ性〉が【重視する観点】として見出された。さらに、入試問題の作成において大切にしていることは、テキスト論の立場から、選ばれた素材文の『肝となる部分をいかにピンポイントに聞くか』であると回答しており、それを〈素材文に対する設問の適切性〉というカテゴリとした。『特に古文漢文の素材文はリソースが限られている』ため〈重複回避〉が課題であり、重複を避けることが『平等性』の担保にもつながると考えていることから、教科書や問題集、他大学を含めた過去問等を徹底的に調査しているという。【問題のタイプ・分類の認識】では、〈現代文の記述問題の分類〉としていくつかのタイプ分類が挙げられ、その分類は難易度とも結び付いて捉えられていた。【測定したい資質・能力】では、〈文章読解力〉と〈答案加工能力〉という国語特有の観点が抽出された。

表3 「作問」におけるカテゴリ分類と回答データの例

中カテゴリ	小カテゴリ	回答例	
素材の収集源	教科書	・教科書の内容に則した問題 (A・物理) ・一番の拠り所は教科書 (C・生物) ・教科書を見てどういふような問題を出しているか (D・数学)	
	本	・さまざまな本を見ながら考えて提案していく (B・化学)	
	他大学の入試問題	・他大学の入試問題を見てどういふような問題を出しているか (D・数学)	
	新聞・雑誌	・欧米英語圏の教養ある人が読む新聞・雑誌 (E・英語)	
重視する観点	難易度	・平均点は意識する、いかにして目的の平均点に近づけるか (A・物理) ・一番簡単な問題でも教科書の例題よりはちょっと難しい問題を意識 (D・数学) ・悪い問題は、誰も解けない問題 (D・数学) ・素材文ごとに難易度を記録している (E・英語) ・過去問を読み込んで、例年から比較した上の難易度のイメージを共有した段階からスタート (F・国語)	
	識別力	・可否の識別性がある問題 (A・物理) ・得点に差がつくように (D・数学)	
	妥当性	・平均点調整って、何を測ってるんだらうってことになる (B・化学) ・回答形式の変更による得点調整はしたくない (B・化学) ・その本文を使うのが高校生に対して適当か適当じゃないか、それによって受験生の力を測定できるのかできないのか (F・国語)	
	重複回避	・他大学の入試問題との被りもダメ (C・生物) ・共通テストと被らないように配慮する (C・生物) ・他大学でも、自大学の問題でも出ていないテーマにする (D・数学) ・全ての教科書と、市販されている問題集と被っているかどうかチェックする (F・国語) ・特に古文漢文はリソースが限られているので、どうしても過去に他大学で出題されたことが出てくる場合もある (F・国語)	
	らしさ	・それじゃあ当該の選抜の種類らしくないので (B・化学) ・工学部には工学部らしい試験問題であってほしいので (B・化学)	
	現実生起するかの確認	・化学の場合、実際にその現象が起きるのか起きないのかが大事 (B・化学)	
	学問観	・ベースになる部分を大事にしなくちゃいけない、生物学っていうのは奥が深いんだよ、ということに気付いてほしい (C・生物)	
	ひっかけ・意地悪な問題の回避	・ひっかけ問題や意地悪な問題は出したいくない (C・生物)	
	教科書・問題集との類似性	・受験生は全然見たこともない問題を出しても、絶対解けない。(D・数学) ・似たようなのをやったことがある問題じゃないと、普通手も足も出ない (D・数学)	
	教材としての価値	・高校の教材として使ってもらえるような問題は良い問題だと思う (D・数学)	
	素材文の語数	・●●ワードから●●ワードくらいまでのもの (E・英語)	
	素材文のテーマの面白さ	・●●大学の問題面白かったなって思っほしい (E・英語) ・なるほどそういうことあるんだっていうふうには、英文を読んで思っほしい (E・英語) ・受験生に対して全く興味関心も感じられない、読め気が失せるものってやっぱりダメだと思う (F・国語)	
	素材文に対する設問の適切性	・このテキストからは、この内容について読み取って、的確に表現してほしいっていうポイントは必ず聞かないといけない (F・国語) ・このテキスト読むんだらう肝はここだろうという部分をいかにピンポイントに聞くか (F・国語)	
	採点のしやすさ	・採点者が採点しやすい問題 (E・英語)	
	入試ミス	・作問において大事にしていることは、事務的なことと言うと入試ミスを出さないということ (E・英語)	
	メッセージ性	・作問において大切にしていることはメッセージ性、何を学んできてほしいかっていうのを明確に打ち出す (B・化学) ・作問において大切にしていることはメッセージ性、見たこともない受験生に対する授業だよな (E・英語) ・それぞれの大学の素材文にはメッセージ性がある (F・国語) ・その素材文を読むと自分が、受験生にとって、少なくとも人生のプラスになるようなものを出さない (F・国語)	
	問題のタイプ・分類の認識	素材 (図)	・物理の場合は、内容的に重なっても素材が違えば全然違う問題 (A・物理) ・物理という素材というのは図、その図が物理の事象を典型的に表している (A・物理)
		記述式問題	・単なる知識じゃなくてそれを組み合わせで情報発信する力、情報発信する力っていうのは説明する力だから、文章を書かせたい (B・化学)
		テーマ	・テーマは素材ではなくて実際に起きていること、大きなカテゴリとして代謝、発生、遺伝、生態、進化、分類がある (C・生物)
		小問形式・大問形式	・普通多くの大学は大問を (1)(2)(3) のように小問に分けるが、大学によっては誘導なし、小問なしという問題もある (D・数学)
語順並べ替え問題		・語順並べ替え問題で文が構成できるっていうのは書く力の要素の1つ (E・英語)	
説明問題		・説明問題はちゃんと英文を読めるかどうかを見ている (E・英語)	
英文和訳問題		・英文和訳は日本語にちゃんとできるかどうかを見ている (E・英語)	
現代文のジャンル		・小説・評論・随想 (F・国語)	
現代文の記述問題の分類	・極端な話、抽象→具体、具体→抽象の2つのベクトルしかない (F・国語) ・一番単純なタイプは圧縮型と呼ばれるタイプで、次は合成型のタイプで、次はピックアップ型のタイプ、一番難しいタイプ具体的な内容しかテキストに書いていない、それを最後自分の言葉を含めて一般化するのが一番難しい (F・国語)		
測定したい 資質・能力	大学が求める資質・能力	・●●大学に来るんだらうこれだけの学力欲しいよねって観点で見ている (B・化学) ・大学がこういう力を求めているんだというのがわかる問題は、良い問題 (E・英語)	
	思考力	・良い問題は、問題をバツと見て答えが浮かばない、解法が浮かばない問題 (A・物理) ・パターン認識で解法が見える問題を解ける学生ではなく、自分で考える力がある学生が欲しい (A・物理) ・知識を活用して考える問題を基本的に出すように努力している (B・化学) ・思考力を問うような問題を出したい (C・生物) ・論理的に考えてやらないと答えが出ないっていうのが良い問題 (C・生物)	
	応用力	・公式を変形して、自分が必要とする形に変えられる力というのはすごく大事 (B・化学)	
	幅広い視野	・幅広い視野、知識を持った人材を育てたい (C・生物)	
	文章読解力	・内容をどこまで掘り下げて理解しているかどうか (F・国語) ・辞書の訳に合ってれば正解ということではなくて、そこに上手い下手とか、本質分かっているのと分かっているの、そこに点数の差がつくのは良い問題 (F・国語)	
	答案加工能力	・字数の制限によって、あんまりちゃんたら長い表現使うと字数超えちゃう、するとそれをシンプルな熟語でまとめないといけないとなってくると、答案加工力が問われる (F・国語)	

表 4 「配点・採点基準の設定と採点」におけるカテゴリ分類

中カテゴリ	小カテゴリ	言及のあった教科・科目
配点の決定方略	回答形式に応じた配点	(B・化学), (C・生物), (E・英語)
	記述・説明問題の配点	(A・物理), (B・化学), (C・生物), (E・英語), (F・国語)
	出題順序に応じた配点	(D・数学)
	識別力に応じた配点	(F・国語)
作問時の難易度調整方略	回答形式の検討	(B・化学)
	問題数の増減	(C・生物)
	問題内容の検討	(D・数学)
	採点基準の明確化	(E・英語)
	素材文のレベルの調整	(E・英語), (F・国語)
設問のレベルの調整	(F・国語)	
実施後の難易度調整方略	配点・採点基準の見直し	(B・化学), (D・数学)

表 5 「試験実施後の評価」におけるカテゴリ分類と回答データの例

中カテゴリ	小カテゴリ	回答例
予想に照らした結果の捉え方	難易度の予想	<ul style="list-style-type: none"> ・幸いにしてほしいの狙い通りにはなっている (A・物理) ・基本難しめになってしまう (B・化学) ・今までやってきた勘で平均点の予想をするが、あんまり外れたことはない (C・生物) ・そんなに難しくないと思って受験生は解けなかつたりする (D・数学) ・難易度の予想は当たらないことが多い (D・数学) ・難易度の予想はほしい当たる、そんなに外れない (F・国語)
	読み違い	<ul style="list-style-type: none"> ・読み違いもあって、こっちができていたと思った問題がことごとくできていない (C・生物) ・こちらの読み違いだよ、受験生はこのへんぐらいできるんだらうと思ったらガタガタ (E・英語)
解釈の観点	識別力	<ul style="list-style-type: none"> ・得点別に可否を分かりやすく示したグラフを作成 (A・物理) ・測定としてはうまく学力のない人とある人を分けた、識別力があつた、と言うんだけど、その識別基準がどこかが重要 (B・化学) ・狙い通りの学力レベルで識別ができていたと感じる問題は、記述がきちんとできていた問題 (B・化学) ・入試だからやっぱり差をつける問題じゃないとダメ (D・数学) ・五分位図による識別力の確認を最も重視 (E・英語) ・読める生徒と読めない生徒を選抜するという意味に関しては、きちんと点数が分かれた問題ではなかったかなと思う (F・国語)
	誤答率・正答率・平均点	<ul style="list-style-type: none"> ・平均点とか、小問別の正答率 (A・物理) ・誤答率がそこそこ高い問題はチャレンジしてくれた問題、解こうっていう気にさせた問題 (B・化学) ・誤答率がめちゃくちゃ低い問題は、簡単すぎた (B・化学) ・誤答率が非常に低く、正答率がやたら高い問題は次年度は問う必要がない (B・化学) ・合格者の平均を重視 (D・数学) ・小問分析で得点率を見ている (E・英語)
	得点の散らばり	<ul style="list-style-type: none"> ・得点の散らばりを見て、得点をとれている人がそこそこいて、でも4割とか3割くらいの人が解けてない時にはうまくいったなって思う (B・化学) ・平均も大事だけど散らばりも大事 (D・数学)
	空欄の比率	<ul style="list-style-type: none"> ・空欄というのは捨てられたか、または解けなかった問題なので、それはどっちかって言うと嫌われた問題 (B・化学)
	記述問題のでき具合	<ul style="list-style-type: none"> ・特に記述のところ、このぐらいなら書けるという部分を見たい (C・生物)
	小問の配点	<ul style="list-style-type: none"> ・得点の解釈では、大問の平均点だけ見ても小問の配点まで見ないとなんとも言えない (D・数学)
	共通テスト基準の難易度把握	<ul style="list-style-type: none"> ・共通テストより難しくなるレベルが難しいんだらうな (E・英語)
	受験者のレスポンス	<ul style="list-style-type: none"> ・できる子は、的確な表現を使って、この表現いいな、っていう解答を作る (F・国語) ・できている子は「おお〜」っていう答えをきちんと揃えてくるので、するとこちらはやっぱり我々の意図したことは間違っていないなと思ひ安心する (F・国語)
入試観	入試の選抜機能	<ul style="list-style-type: none"> ・選抜試験だから、選抜に適した識別性がないと意味がない (A・物理) ・定員を確保するための試験と、定員はどうでもいから能力を確保したいという試験は違うので、大学入試はうまく定員のところでギャップが出るような問題にしないでほしいところが難しい (B・化学) ・入試問題だから、入試の選抜に役立つ資料として識別力が一番大事 (E・英語)
	入試の教育機能	<ul style="list-style-type: none"> ・高校に負担をかけないで、かつ、こういうところは大事よみたいなメッセージが送られれば (A・物理) ・初年度は受験生に厳しくても、その質とレベルを保てば、5,6年経てばそれに対応した受験生が現れてくるので、受験生に対してこちらが求めている能力を持っている、またそれを養ってきた受験生が受けてくれるようになる、と言うか、受ける受験生はそれを養ってから来るようになる、それはすごく教育効果があると思う (B・化学) ・続けて出していると、対策してきたらできるようになるかなと思つてね、そしたらそれはそれでいいことだから、教育効果があるわけだから (E・英語) ・大学入試の問題を解くことによって、有益な読書となるような (F・国語)
	下級学校への悪影響の排除	<ul style="list-style-type: none"> ・作題において大切にしていることは、高校の授業に悪い影響を与えないということ (A・物理)
	高校教育へ思量	<ul style="list-style-type: none"> ・問題のでき具合から考えると、高校で実験とかやってないのかなって (C・生物)
	将来への配慮	<ul style="list-style-type: none"> ・入試だけじゃなくてその後の大学生活とか、あるいは社会人になった時に、考えるってことが生きてくる (C・生物)

3.2.2 「配点・採点基準の設定と採点」のフェーズ

【配点の決定方略】の〈回答形式に応じた配点〉は、回答形式によって測定できる資質・能力の種類や深度に差があるという認識が背景にあるものと考えられる。

〈記述・説明問題の配点〉は、ほとんどの教員から言及があったカテゴリであり、記述式問題は多くの作題者にとって重要な要素と認識されていることが示唆された。教員 D の発話からは〈出題順序に応じた配点〉というカテゴリを生成したが、これは数学の問題固有の特性に依拠する可能性がある。また、教員 F は『識別力の高い問題は配点が高くあるべき』という信念を持っていた。

【作問時の難易度調整方略】としては、〈回答形式の検討〉〈問題数の増減〉〈問題内容の検討〉〈採点基準の明確化〉〈素材文のレベルの調整〉〈設問のレベルの調整〉という 6 つの小カテゴリが抽出された。それぞれ、各教員の専門分野ごとの問題形式の特徴を反映する方略であると解釈できる。

教員 B と教員 D は、【実施後の難易度調整方略】に言及していた。ただし、いずれも【実施後の難易度の調整方略】を肯定的に捉えた内容ではなく、〈配点・採点基準の見直し〉等により試験実施後に得点調整を行うと、『受験生の力を正確に把握できないのではないか』『本当の力が測れていないのではないか』といった、妥当性や識別力の毀損への懸念を意味するものであった。

3.2.3 「試験実施後の評価」のフェーズ

実施後の得点率から見た難易度が作問時の予想に比して当たっているか否かを振り返った〈難易度の予想〉や、できるはずだと思って出題した問題ができていなかったという〈読み違い〉に関する認識を【予想に照らした結果の捉え方】とした。教員 B や教員 D は予想よりも実際の得点率が低くなることが多く、予想は当たらないと述べていたのに対し、教員 A・C・F は、これまでの経験に基づく平均点予想は当たる、と述べていた。

実施結果に対する【解釈の観点】の 1 つとして〈識別力〉が見出された。ただし、『五分位図による識別の確認を最も重視する』という教員 E のように、統計的な観点からエビデンスに基づく評価を意図しているケースから、『うまく学力のない人とある人を分けた』『差をつける問題でないといけない』という概念的理解にとどまるケースまで、識別力に関する認識や理解のレベルには個人差が見られた。また、〈得点の散らばり〉というカテゴリは、〈識別力〉と結びついて解釈されて

いるものと推測される。

教員 B は、〈誤答率・正答率・平均点〉や〈空欄の比率〉に独自の意味付けをして、結果の解釈を行っていることが示された。教員 D は、受験者全体の得点率よりも『合格者の平均』を重視していた。また、「作問」のフェーズで【測定したい資質・能力】に〈思考力〉〈幅広い視野〉を挙げている教員 C は、「試験実施後の評価」のフェーズで重視する観点として〈記述問題のでき具合〉に言及していた。これは、記述問題によって〈思考力〉や〈幅広い視野〉が測定できるという認識に基づいているものと考えられる。国語を専門とする教員 F の【解釈の観点】としては、〈受験者のレスポンス〉が重要な役割を果たしていることが示された。

「試験実施後の評価」のフェーズで抽出されたカテゴリの中でも、個別大学の入試問題の機能や在り方に関する信念を反映したカテゴリをまとめて【入試観】とした。教員 A・B・E からは、〈入試の選抜機能〉に注目した発話が見られた。〈入試の教育機能〉では、教員 F は入試問題を解くことも教育の一環であるべきという受験時の一時点における教育機能を指している一方で、教員 B や教員 E の発話内容は、入試の持つ波及効果 (washback effect) (村山, 2006) を意図した概念であり、高校教育における試験対策を是とする考え方が示唆されている。これに対して、教員 C の〈高校教育への思量〉と〈将来への配慮〉は、過去に遡って高校教育の現状を把握するための一種の鏡として入試を捉える見方と、未来へ目を向けて入学後、あるいは社会に出てからも必要とされるようなポテンシャルを評価すべきであるという信念の両面を併せ持つ入試観を表している。また、教員 A からは、いわゆる 46 答申の中間報告で示された「日本型大学入学者選抜制度の三原則」(木村・倉元, 2006) の 1 つである〈下級学校への悪影響の排除〉に相当する強い思いが語られた。

3.3 作題関連業務の規定要因に関する仮説モデル

作題関連業務の 3 つのフェーズでは、教員の専門が違っても共通する小カテゴリが数多く生成された一方で、各専門分野に特有の学問観、個人の教育観等を反映していると推察される小カテゴリも複数抽出された。また、業務の手順に応じて、便宜的に 3 つのフェーズに分けて大カテゴリとしたが、実際にはそれぞれのフェーズは有機的に関連しているものと考えられる。そこで、抽出されたカテゴリをもとに、各フェーズを規定する要因を整理し、作題関連業務の背後にある認知プロセスに関する仮説モデルを構築して図 1 に示した。

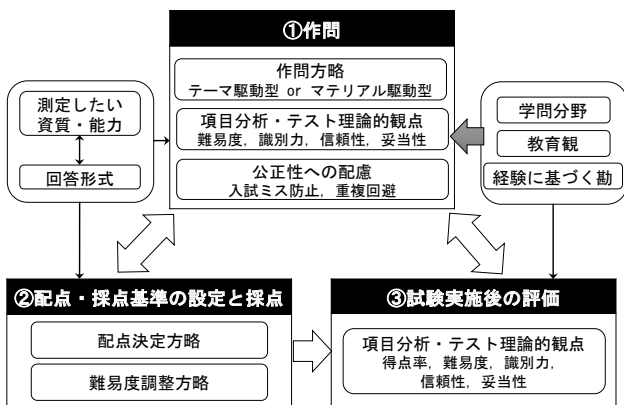


図 1 作題関連業務の背後に仮定される認知プロセス

まず、面接調査に基づくカテゴリ生成の結果より、①「作問」と②「配点・採点基準の設定と採点」、そして①と③「試験実施後の評価」は、単一年度のみならず年度を跨いだ作題関連業務の一連のサイクルの中で相互に関連していることが示唆された。一方で、②の採点結果を踏まえて③の評価は行われるが、③から②への遡及的な矢印は事後的な難易度調整を暗に意味するため、妥当性保証の観点からむしろ避けることが望まれていた。

①「作問」では、テーマ駆動型かマテリアル駆動型かといった「作問方略」、テストの性能を評価する観点である「項目分析・テスト理論的観点」、そして入試としての品質に関わる受験者の「公正性への配慮」という3つの要因が重要視されていると考えられる。ここで、「作問方略」におけるテーマ駆動型は主に理科を専門とする教員 B・Cに見られた、まずはテーマ、すなわち問題で取り上げる領域等の主題から決定するという方略を指す。これに対して、マテリアル駆動型の典型は、英語および国語を専門とする教員 E・Fによって語られたような、素材文を最初に決定するという方略を指す。この作題方略の違いには「学問分野」の影響が大きいと推察される。数学はどちらかと言えばテーマ駆動型に相当し、物理は、テーマ駆動型とマテリアル駆動型の双方の要素を併せ持った学問分野であると推測される。

作題関連業務に従事する教員には、「測定したい資質・能力」とそれを測定し得る「回答形式」が対応しているという認識を持っていることが示唆された。この認識は難易度や識別力に関する作問時の見通しを規定しており、さらに②「配点・採点基準の設定と採点」のフェーズにおいても、難易度は回答・問題形式によって調整可能であり、識別力は配点によって調整可能であるという認識を生み出していた。

また、作題関連業務に従事する教員は、何らかの「項目分析・テスト理論的観点」に基づき③「試験実施後の評価」を行っていることが示された。しかしながら、科学的なテスト理論への理解度には個人差が大きく、経験に基づく勘や各人の教育観に依拠した解釈も見受けられた。さらに、難易度と識別力との混同や、識別力と配点との関係性に関する誤謬なども散見され、作題者の多くがテスト理論の正しい知識を有していないことが露呈された。

以上より、作題関連業務に携わる教員へのFD等を通してテスト理論の知識と方法を教授する、情動的支援が有効である可能性が示された。また、公正性保持のための要件となる入試ミスの防止や重複回避に関しては、高校教科書の範囲や用語、自大学および他大学の過去問等をデータベースとして参照するようなチェックシステムによる道具的支援が有用であろう。

4 本研究の限界と今後の展望

本研究では、面接調査に基づく質的な分析により、作題関連業務を規定する要因を探索的に抽出し、図1のような認知プロセスに関する仮説を生成した。現状では、学問分野による違いに加えて、個人のパーソナリティや経験に依拠した信念も一定の影響を持つ要因であることが示唆された。同時に、作題支援の具体的な方策についても明確化できた。作題関連業務に従事する教員に直接アプローチすることで、認知プロセスを解明する試みの有効性を示唆する結果であり、パイロットスタディとしての意義は大きい。

しかしながら、調査対象者が6名と非常に少なく、物理、化学、生物、数学、英語、国語の教科・科目から各1名のみであったため、学問分野固有の特質と、個人の特性とを区別して論じることができない点が、本研究の限界として挙げられる。今後は、調査対象者の人数を増やし、幅広い専門からの知見を収集することで、教科・科目特異性の高い要因、あるいは教科・科目横断的な汎用性の高い要因を特定するなど、仮説をさらに精緻化し、量的な分析手法の適用によって、その仮説を検証していく予定である。これらの研究成果の蓄積をもとに、個別大学における作題関連業務の効率化を図るとともに、良質な入試問題の継続的作成に資する作題支援システムの構築を目指したい。

注

1) 本稿は、令和3年度全国大学入学者選抜研究連絡協議会大会(第16回)で発表した内容(久保, 2021)に、新たなデータを加えて再構成したものである。

- 2) 一義的に解釈できる意味内容ごとに断片化した文書の1つのまとまりを、1つのデータとして扱っている。1つのデータは、一文の場合もあれば、複数の文から構成される場合もある。

謝辞

本研究は、東北大学 2020 年度 TUMUG 支援事業 (男女共同参画・女性研究者支援事業)「スタートアップ研究費」3-20 および、JSPS 科研費 JP21K13636 の助成を受けた研究成果である。

参考文献

- 木村拓也 (2012). 「個別学力検査問題の質を測る—作題に関する合格者アンケートによる試み—」『大学入試研究ジャーナル』 **22**, 135–144.
- 木村拓也・倉元直樹 (2006). 「戦後大学入学者選抜制度の変遷と東北大学の AO 入試」『東北大学高等教育開発推進センター紀要』 **1**, 15–27.
- 久保沙織 (2021). 「個別大学の入試問題作成において重視される観点の探索的分析—作題支援システムの構築を目指して—」独立行政法人大学入試センター『令和 3 年度全国大学入学者選抜研究連絡協議会大会 (第 16 回) 研究発表予稿集 (クロードセッション用)』, 62–67.
- 倉元直樹 (2020). 「受験生保護の大原則と大学入試の諸原則」倉元直樹編『「大学入試学」の誕生』金子書房, 6–17.
- 村山航 (2006). 「教育評価」鹿毛雅治編『朝倉心理学講座 8 教育心理学』朝倉書店, 173–194.
- 佐藤喜一・並川努・濱口哲 (2013). 「新潟大学の個別学力検査における作題技術の向上を目的とした作題者へのフィードバック内容の整理」『大学入試研究ジャーナル』 **23**, 127–133.