

テスト問題作成ガイドラインの開発（1）

— 日本での普及に向けた整理 —

石井秀宗¹、荒井清佳²、坪田彩乃¹、安永和央³、寺尾尚大²

¹名古屋大学大学院教育発達科学研究科，²大学入試センター研究開発部，

³IPU・環太平洋大学次世代教育学部

問題と目的

入試や資格試験をはじめ、教育の様々な場面でテストは用いられているが、日本においては、テスト作成に関する知見はあまり蓄積されていない。英語など個別の教科において、テスト作成について解説する書籍はいくつかあるが（笠原・佐藤, 2017 など）、問題作成やテスト開発について専門的に扱う文献や書物は僅かしかないのが現状である。この背景には、日本では教員養成課程において教育測定学やテスト理論が教えられておらず（木村, 2010）、多くの場合、教員が経験を頼りに試験問題を作っていること、また、入試のようなハイ・ステークスなテストでさえも、テストの設計、問題作成、編集等を、当該学問領域の教員が兼業として行っていることなどの日本のテスト文化がある（石井, 2018）。

これに対し諸外国では、教育測定学やテスト理論が研究領域として広く認知され、テスト問題の作成等に関する学術書も多数出版されている（Haladyna & Rodriguez, 2013 など）。それらの中では問題作成ガイドラインも示され、**Item Writer** や教員等がテスト問題を作成する際に利用できるようになっている。

日本においてもこのようなガイドラインが必要であるが、テストの構成や文化の違いなどから、海外のものをそのまま持ち込んでも十分には機能しないと考えられる。そこで本研究では、日本においてテスト問題作成ガイドラインが普及することを目指し、その手始めとして、海外のガイドラインや、これまでの研究で得られている知見（荒井, 2015; 坪田・石井, 2020; Terao, 2020; 安永・齋藤・石井, 2012 など）を参照して、日本のテスト事情にあったテスト問題作成ガイドラインの整理を行う。

ガイドライン開発の基本的な考え方

日本においては職業的な **Item Writer** はほとんど存在せず、学校教員など当該領域の専門家がテスト問題を作成している。そこでガイドラインの開発にあたっては、学校教員等を主なユーザーとして念頭に置くこととする。

先にも述べた通り、日本の教員養成課程においては教育測定学やテスト理論等が教えられていないことから、それらの領域の専門用語を説明なく用いることのないようにする。

教員が作成するテストは主としてクラスルームテストであることから、対象とするテストは主にクラスルームテスト、小学校であれば教師作成テスト、中高であれば定期テスト、大学等であれば学期末試験等を念頭に置く。また、入試などハイ・ステークスな大規模テストも重要であることから、それらの問題作成にも資するガイドラインを目指す。

問題形式としては、クラスルームテストや大規模テストでよく用いられる、多枝選択式問題、空所補充問題、記述式問題を主に想定する。

各ガイドライン項目は単体で理解できるようにする。抽象的な一般論ではなく、具体的にどのようにしたら良いかがイメージできることを目指すが、状況やユーザーによって解釈に幅があっても構わないところは細かく規定せず、適度なところを押さえるようにする。

ガイドライン項目はなるべく簡潔で明快なものにする。将来的には、すべてのガイドライン項目に説明を付し、より詳しくは説明を読んで理解するような構成とする。

ガイドライン項目の説明

Haladyna & Rodriguez (2013) やこれまでの研究等を参照して整理したテスト問題作成ガイドラインを表1に示す。ここでは、いくつかの項目について説明を付す。

9. 「あてはまるものをすべて選べ」という設問は避けること。使う場合は部分点を与えること

この形式の設問では、正しいもののみをすべて選んだ場合を正答、それ以外を誤答とすることが多い。しかし、そうすると例えば、判断が難しい選択枝が1つでもあると正答率が極端に低くなってしまったり、またそのような危険性から受検生の不安が高まり、十分に能力を発揮できなくなってしまうなどの問題がある。これらの問題は部分点を与えることである程度回避できる。部分点の与え方としては、荒井・宮埜 (2017) などの検討もあるが、簡潔な方法は、部分点を与える解答（例えば「正しいものを1つだけ選び損ねている場合」など）をあらかじめ決めておき、正答の場合の得点の半分を与えるといった方法が考えられる。

10. 読解力や思考力を測る記述式問題では字数制限を設けないこと

読解力・思考力と、制限字数（多くの場合、所定の9割程度の文字数）で文をまとめる力は異なる能力である。字数制限を設けた場合、内容は読み取れているが字数に収められないため不十分な解答になることがあり、得点率や識別力が低下してしまう（安永・齋藤・石井, 2012）。これは思考力を問う場合にも生じる問題であり、読解力や思考力を問いたい問題では字数制限を設けるべきでない。記述量のおおよその目安を示したい場合は、解答欄の大きさを調節することで可能となる。

16. 空所補充問題について、文意が分からなくなるほどの空所を設けないこと

(1) を合計して (2) で割った値を (3) 値と言う。(1) を大きさ順に並べたとき、ちょうど (4) にくる値を (4) 値と言う。(2) が (5) のときは、(4) の2つの (1) の (3) を (4) 値とする。(2) が最も大きい (1) の値を (6) 値と言う。

例えば、上記は教科書の文章をほぼそのまま用いて空所補充問題にしたものである。答えを知っていれば、つまり、教科書を丸暗記していれば簡単な問題だが、答えを知らなければ、解答の手掛かりとなる情報が不足していて、パズルを解くような問題になっている。しかも、内容を学習・理解できていれば良いのではなく、もとの文を再現するように考えて解答することが求められるので、知識問題としても内容理解問題としても適切でないものになっている。補充すべき語句を重要なものだけに絞り、内容を理解していれば正答できるように、手掛かりとなる情報は残すように空所を修正するか、問題形式を変えるなどの対応が求められる。

24. 不必要に選択枝を増やさないこと

多枝選択式問題において、多くの場合、選択率がある程度大きい選択枝数は3枝程度なので、あて推量による正答の可能性を低くするためとして選択枝を多く設定するのは無意味である。出題ミスやマークミスの原因にもなるので、不必要に選択枝を増やすのは避けたほうが良い。実際のテストでは4~5枝が設定されることが多いが、4つめ以降の選択枝を作るのが困難であったり、多大な労力を要するような場合は、3枝でも十分である。

複数の空所に対して、共通選択枝群の中からあてはまるものをそれぞれ選ぶ形式の問題においては、共通選択枝の数は、受検者が選択枝を一覧するのに困難を伴わない程度が望ましい。選択枝が多くなると、選択枝を探すのに時間を費してしまうだけでなく、出題ミスやマークミスなどの問題も生じる。

今後の課題

本研究では、クラスルームテストや大規模テストで用いられることの多い、多枝選択式問題、空所補充問題、記述式問題を主に想定してガイドラインを整理したが、他の形式の問題も教育現場等では利用されている。それらの問題についてのガイドラインも今後検討し、整備する必要がある。昨今普及が進むCBTでは、実施や問題提示、解答方法など、紙筆式テストとは大きく異なる部分があることから、CBTに対応するガイドラインの策定も急務であると考えられる。また、荒井(2015)のように、実際にテストを作成している教員等からのフィードバックを得て、ガイドラインを修正していくことも必要である。

文献

- 荒井清佳(2015). 多枝選択式問題を作成する上で大切なこと—問題作成の専門家に対する調査結果に基づいて—. 日本テスト学会誌, 11, 21-34.
- 荒井清佳・宮埜寿夫(2017). 正答選択肢をすべて挙げる形式の問題の採点について. データ分析の理論と応用. 6, 101-112.
- Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. Routledge.
- 石井秀宗(2018). 大学入試における共通テストの複数回実施は実現可能か—日本のテスト文化やこれまで見送られてきた理由などからの検討—. 名古屋高等教育研究, 18, 23-38.
<https://www.cshe.nagoya-u.ac.jp/publications/journal/eighteenth.html>
- 笠原 究・佐藤臨太郎(2017). 英語テスト作成入門—効果的なテストで授業を変える!— 金星堂
- 木村拓也(2010). 日本における『テストの専門家』を巡る人材養成状況の量的把握. 日本テスト学会誌, 6, 29-49.
- Terao, T. (2020). Difficulty of summarization items for Japanese learners: Effects of passages, distractors, and response formats. *Frontiers in Education: Assessment, Testing and Applied Measurement*, 5, 1-15. <https://doi.org/10.3389/feduc.2020.00009>
- 坪田彩乃・石井秀宗(2020). 多枝選択式問題作成ガイドラインの実証的検討. 日本テスト学会誌, 16, 1-12.
- 安永和央・齋藤 信・石井秀宗(2012). 構造的性質を操作した国語テストにおける回答の検討—中学生を対象にしたテストの実証研究—. 日本テスト学会誌, 8, 117-132.

表1 テスト問題作成ガイドライン (2021.9)

<p>問題の内容</p> <ol style="list-style-type: none"> 1 問いたいことは何か, 問題を解くために必要な能力は何かが明確であること 2 重要な事柄を問うこと。些末なことや一般的過ぎる問いにならないこと 3 正解が問題作成者の価値観に左右されるような問いにならないこと 4 特定の個人や集団に有利または不利な内容にならないこと 5 ひっかけ問題にならないこと 6 高次の能力を測る問題では, 受検者にとって新奇な素材を用いること <p>問題の形式</p> <ol style="list-style-type: none"> 7 測りたい能力に見合った問題形式を用いること 8 前の問題に対する解答が, 後の問題の正誤に影響しないこと 9 「あてはまるものをすべて選べ」という設問は避けること。使う場合は部分点を与えること 10 読解力や思考力を測る記述式問題では字数制限を設けないこと <p>問題の記述</p> <ol style="list-style-type: none"> 11 言語レベルを受検者集団に合わせること 12 教示文・本文・設問・選択枝・図表等の記述量を必要最小限にすること 13 教示文・本文・設問・選択枝・図表等の文言をよく校正すること。他の人にも確認してもらうのが望ましい 14 教示文・本文・設問・選択枝・図表・解答欄等のレイアウトや大きさを適切にすること 15 とくに低学年の児童に対して, 選択枝は行を変えて1つずつ並べること 16 空所補充問題について, 文意が分からなくなるほどの空所を設けないこと <p>設問部分</p> <ol style="list-style-type: none"> 17 問いたいことは何か, どのような形式で解答したらよいかを明確・簡潔に書くこと 18 本文や選択枝など他の部分を読まなくても, 設問部分だけで何を問われているかが分かること 19 否定表現を使わないこと。もし使う場合は, 太字や<u>アンダーライン</u>で強調すること 20 一部の受検者にしか分からないような暗黙の前提を用いないこと <p>選択枝</p> <ol style="list-style-type: none"> 21 いずれの選択枝ももってもらいたいこと 22 高得点者と低得点者をよく区別できるような, 識別力の高い選択枝を用いること 23 正答枝と誤答枝が明確に区別できること 24 不必要に選択枝を増やさないこと 25 明らかな誤答枝やお遊びの選択枝など, 余計な選択枝を入れないこと 26 五十音順, 数量の大きさ順など, 何らかの法則に従って選択枝を並べること 27 正答枝の位置をランダムにばらつかせること 28 「上記のいずれでもない」「上記すべてあてはまる」などの選択枝を用いないこと 29 「～でない」「～以外である」など否定表現を用いないこと 30 「絶対に」「常に」「決して」「完全に」など, 強意語を用いないこと 31 選択枝は互いに独立であること。内容に重なりがないこと 32 一方が正答枝であれば他方は誤答枝であると分かるような, 両立しない選択枝を入れないこと 33 選択枝の長さをおおむね揃えること 34 選択枝の内容や形式などの構造を揃えること <p>記述式問題の採点に関して</p> <ol style="list-style-type: none"> 35 問題作成と同時に, 評価の観点, 採点基準を設定すること 36 評価の対象とする要素, しない要素を明確にすること (誤表記は減点する, 乱筆は減点しない, など) 37 正答となるものについて, 基本的な基準を設定し, あまり細かな条件設定はしないこと 38 本格的に採点を始める前に, いくつかの解答を採点して, 採点基準を精錬すること 39 最初のほうで採点した解答を後で採点し直すなどして, 評定の一貫性を保つこと
--