

難易度と識別指数を用いた単文統合型作問課題の分析

Analysis for Problem-Posing Assignment in MONSAKUN Based upon Difficulty and Discrimination Index

岩井 健吾^{*1}, 松本 慎平^{*2}, 林 雄介^{*3}, 平嶋 宗^{*3}

Kengo IWAI^{*1}, Shimpei MATSUMOTO^{*2},
Yusuke HAYASHI^{*3}, Tsukasa HIRASHIMA^{*3}

^{*1} 山陽女子短期大学人間生活学科

^{*1} Department of Human Life Studies, Sanyo Women's College

Email: iwai@sanyo.ac.jp

^{*2} 広島工業大学情報学部

^{*2} Faculty of Applied Information Science, Hiroshima Institute of Technology

Email: s.matsumoto.gk@cc.it-hiroshima.ac.jp

^{*3} 広島大学大学院工学研究科

^{*3} Graduate School of Advanced Science and Engineering, Hiroshima University

Email: {hayashi, tsukasa}@lel.hiroshima-u.ac.jp

あらまし：和差の算数文章題を対象とした単文統合型作問学習支援環境「モンサクン」が設計・開発されている。このモンサクンの作問課題は、小学生から大人まで幅広い層の利用実績があり、演習全体の学習効果も確認されている。その一方で、個々の作問課題の適切性に関しては、学習者の演習データに基づいて実施されていない現状にある。本研究では、古典的テスト理論の指標である難易度と識別力の観点から、過去の実践利用で収集した学習者の演習データを用いて個々の作問課題の適切性の分析を実施した。

キーワード：作問課題，正答進行型課題，正答率，識別指数

1. はじめに

和差の算数文章題を対象とした単文統合型作問学習環境「モンサクン」が設計・開発されている⁽¹⁾。

このモンサクンの作問課題は、既に小学生から大人まで幅広く利用実績があり、演習全体を通じた学習効果も既に確認されている。また、算数文章題の構造的記述に基づくモデルに基づいた個々の作問課題の適切性に関する分析も実施されている。学習課題のモデルに基づいた分析によって、課題自体がもつ客観的な性質を明らかにすることが可能となる。

その一方で、学習者の演習データに基づいた個々の作問課題の適切性に関しては十分に分析が実施されていない現状にある。学習課題のモデルに基づいた分析によって作問課題の自体の性質を明らかにすることは可能であるが、その作問課題が学習者に適したものとなっているかどうかは別途データを用いて評価を行う必要がある。

課題の適切性を評価する方法として、古典的テスト理論の難易度と識別力に関する議論がある⁽²⁾。この指標に基づいて分析を行うことで個々の作問課題が学習者にとって適切であったかどうかを判断することが可能となる。

そこで、本研究では、古典的テスト理論の指標である難易度と識別力に基づき、過去の実践利用で収集したモンサクンの演習データを使用して、個々の作問課題の適切性の評価を実施した。その結果、1) 出題された作問課題は学習者にとって概ね適切な課題設定であったこと、2) 同レベルに異なる性質をもつ作問課題が含まれていたこと、が明らかとなった。

2. 単文統合型作問支援環境「モンサクン」

与えられた単文を組み立てることによって、算数文章題の問題を作成する単文統合型作問支援環境「モンサクン」が設計・開発されている。図1は、モンサクンの演習画面の外観である。画面左上側に作成すべき問題の条件(問題制約)が提示され、画面右側に問題の構成要素である単文カードが提示されている。この画面において、学習者は、図1の画面右側に与えられた単文カードを画面左下の3つの空欄に問題制約を満たすように割り当てることで問題の作成を行う。問題作成後、画面左側の一番下にある診断ボタンを押すことで正誤診断が支援環境上で実施され、その結果が学習者にフィードバックされる。正答の場合は次の課題に進み、誤答の場合は、正答に到達するまで何度も回答を繰り返す課題(正答進行型課題)として設定されたものとなっている。

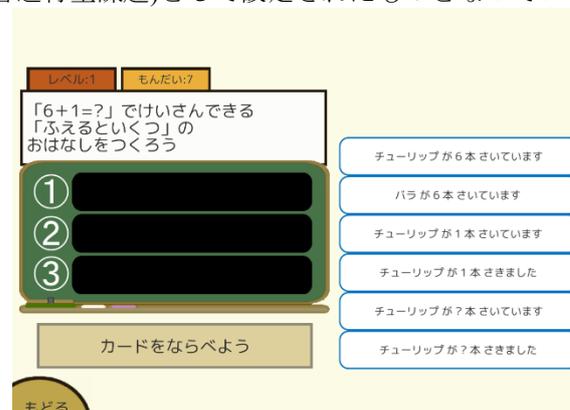


図1 モンサクンの演習画面の外観

3. 分析方法

本研究では、難易度と識別力の観点から、過去の実践利用で収集したモンサクンの演習データを用いて個々の作問課題の適切性の分析を実施した。

3.1 演習データの収集・選定

2020 年度に公立小学校の全学年全クラスにおいてモンサクンの授業内利用が行われた。本研究では、その際に収集された演習データを対象に分析を実施した。この実践では、3 つのレベルが出題され、レベル 1：16 課題(4 課題はチュートリアル用の課題)、レベル 2：12 課題、レベル 3：20 課題、の構成であった。本分析では、レベル 2 を全課題演習した被験者(495 名)に絞って分析した。また、レベル 1 のチュートリアル用の課題は分析対象から除外した。なお、村上の先行研究では、被験者の絞り込みをせずに分析を実施している⁽³⁾。また、レベル 3 を全課題演習した被験者に絞った分析、各学年に絞った分析を実施した点も本分析の新規性であるが、この 2 つの分析方法・結果に関しては当日報告を行うものとする。

3.2 データ分析の手順

難易度の指標に正答率、識別力の指標に識別指数を用いて分析を行った。正答率は、各課題の正答者数を被験者数で割って算出した。識別指数は、上位群 25%、中位群 50%、下位群 25%の 3 群に分類した。順位付けは、各課題の正答を 1 点、誤答を 0 点とした、レベル 1 と 2 の総合成績(28 点満点)を使用した。

3.3 得点化の手順

本分析では、得点化の手順が通常とは異なり、複数の採点基準を設定している。モンサクンの作問課題は、正答進行型課題であるため、指定した回答回数以下の場合を得点化時の正答とした。この正答の判定基準の閾値は、人が恣意的に決めることは妥当ではない。それゆえ、本分析では、その閾値が 1 回以下~20 回以下に設定して得点化を実施した。そのため、各閾値に正答率と識別指数の値が存在するものとなる(各課題 20 パターン存在する)。それらの候補の中で識別指数が最大となる採点基準の閾値を最終的な採点基準の閾値として採用し、正答率と識別指数を決定した。

4. 分析結果

前章で述べたように本分析においては、識別指数が最大となる採点基準の閾値を採用した。図 2 に各課題で最大となる採点基準の閾値を示す。レベル 1 の課題は、12 課題中 9 課題が 1 回以下を正答とする場合に識別指数が最大となった。7 番目は 3 回以下、15 番目と 16 番目は 2 回以下の場合に最大となった。レベル 2 の課題は、12 課題中 10 課題が 1 回以下を正答とする場合に識別指数が最大となった。2 番目は 3 回以下、10 番目は 2 回以下の場合に最大となった。図 3 に、図 2 の採点基準の閾値に対応する各課題の正答率と識別指数を示す。全ての課題の識別指数の値が、高い識別力とされる 0.25 以上であった。

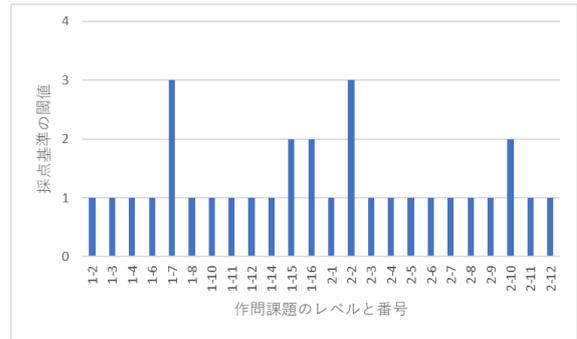


図 2 各課題の採点基準の閾値

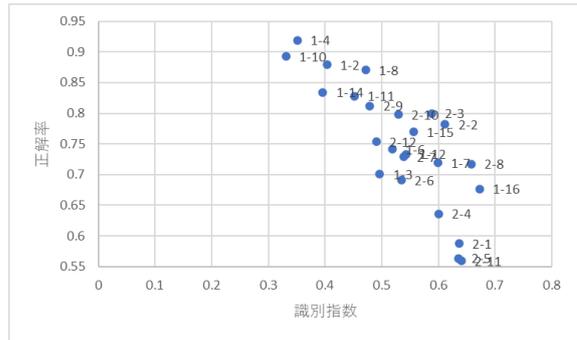


図 3 各課題の正答率と識別指数

5. 考察

図 3 の結果から、難易度および識別力には問題がなかったといえる。図 2 の結果から、同レベルであっても識別指数が最大となる採点基準の閾値が異なることが明らかとなった。これは、同レベルに属する別の課題とは性質が異なる可能性を示唆している。これらの課題に対してさらに学習課題のモデルに基づいて、詳細な分析を行った結果、別の課題とは異なる性質を持つことが明らかとなった。

6. まとめ

本研究では、モンサクンを対象に難易度と識別力の観点から演習データに基づいた作問課題の適切性の分析を行った。その結果、1) 作問課題は概ね妥当であったこと、2) 同レベルに性質の異なる作問課題が含まれることが明らかとなった。

参考文献

- (1) 山元翔, 神戸健寛, 吉田裕太, 前田一誠, 平嶋宗: “教室授業との融合を目的とした単文統合型作問学習支援システムモンサクン Touch の開発と実践利用”, 電子情報通信学会論文誌, 第 J96-D 巻, 第 10 号, pp. 2440-2451 (2013)
- (2) 安永和央, 石井秀宗: “テストにおける設問の問い方が回答傾向に及ぼす影響—国語読解テストを用いた実証研究—”, 教育心理学研究, 第 60 巻, 3 号, pp. 296-309 (2012)
- (3) 村上大希, 松本慎平, 岩井健吾, 林雄介, 平嶋宗: “作問学習システム「モンサクン」における学習ログデータ分析—識別指数を用いた学習課題の評価—”, 中国地区 教育システム情報学会 2021 年度学生研究発表会, pp.149-150(2021)