

研究指導 中澤 真 教授

プログラミング教育における作問学習による学習効果向上の定量的評価

—プログラミング的思考の観点から—

佐藤 晴日

1. はじめに

近年、人工知能の発展やIoT¹技術の進展から、「第四次産業革命」と呼ばれる時代の到来が注目されており、それによる人間の活動の活性化が期待されている[1]. この社会の動きに関して、情報化社会で生き抜くために必要となる資質・能力を学校教育の中で育成するために、小学校段階でのプログラミング教育の導入が注目されている[2]. 特に、プログラミング教育では文部科学省が教育において重要だとしている思考力[3]も強化することができるため[4], 様々な視点から問題を分析して問題点を発見し、解決に導く問題解決力・問題発見力の向上も期待されている.

一方、思考力を養い、理解の深化を図る手法として作問学習が古くから取り組まれている[5][6][7]. このため、相乗効果をねらってプログラミング教育に作問学習を導入することも試みられている. 実証研究もいくつか実施されており、その効果についても受講者の最終的な成績が向上したということが報告されている[8][9]. 具体的には、学習者が作成した問題文に対する評価点と対象単元の間試験の点数の間に弱いプラスの相関があることを示されている. しかし、相関係数がかかなり小さいことに加え、学習者が作成した問題文の評価点を決める基準が厳格でないため、学習効果の定量的な評価としては不十分な点がある.

そこで、本研究では作問課題実施によるプログラミング教育における学習効果向上を単元理解の深化という観点に基づいて検証する. そのためにまず、作問課題のためのルーブリック²を作成し、学習者が作成する問題文の評価基準を明確化する. 最終的には、ルーブリックに基づく作問の評価点と、対象単元の小テストや練習問題を用いた学習者の理解状況との関係について定量的に分析してこれを明らかにする.

2. プログラミング教育における作問学習

2.1 プログラミング教育の現状・効果

近年では情報化社会の進展による人工知能の発展から、人工知能に人間の職が代替されることが危惧されている[1]. そのため、これからの社会において、

情報化社会の中で生き抜くために必要となる能力を学校教育の中で育成することが重要になる. 情報化社会で求められている能力の一つにプログラミング的思考がある. これは、自分が解決しようとする問題について、その解決に必要な動作の組み合わせについて考えるというものである. また、その考えた内容が間違っていた場合にはそれを自分で修正するなどして、課題を解決するために必要となる物事について論理的に考えていく力のことでもある[10]. この力を修得するために小学校段階からのプログラミング教育が重要と考えられ、2020年度から必修化されることになっている[2]. さらに、プログラミング的思考の鍛錬を通して、文部科学省が重要だと考えている思考力[3]の強化につながることも期待されている.

2.2 作問学習の概要

作問学習は学習者の思考力を養い、学習に関する理解の深化を図るための手法であり、その学習効果についても広く知られている[5][6][7].

作問学習とは、現在学習している単元などを題材とした問題を、教員ではなく学習者が作成する活動のことを指し、学習者が主体的に取り組む学習方法である. 従来の学習方法は、教員から学習者に対して一方的に復習問題などが提示されるという受動的なものである[11]. これに対して作問学習では、学習者が主体となることから、従来の学習よりも効果的な復習が可能になるとされ、単元理解の定着に役立つとされる[12]. そのため、同じように思考力や理解度について効果があるとされるプログラミング教育に、作問学習を導入することで、これらに対する相乗効果が期待できる.

2.3 プログラミング教育における作問学習の現状

プログラミング教育において作問学習を実施し、その効果について検証した研究は既にある. 山岸[8]は、学習者が作成した問題文を教員が評価した点数と、作問学習の対象単元についての中試験の点数との相関を算出して作問学習の効果を分析している.

算出された相関は、+0.109と、相関係数がかかなり小さくなっていることから、単元理解の向上について十分な結果が出ているとはいえない. また、問題文の評価については工夫度しかみておらず、その評価基準が曖昧な部分がある.

¹Internet of Things の略

²学習到達度を示す評価基準を表として示したもの

別の取り組みとしては、学習者の相互評価を軸にした作問学習を実施し、アンケート調査による学習者の意識の変容と成績とを比較して学習効果を分析したものがあ[9]。この研究では、学習者が作成した問題文の出来不出来については評価せず、アンケート調査のみで作問学習がプログラミングの知識・理解度の向上に役立つと結論付けている。

以上のことから、既存の研究ではプログラミング教育における作問学習の効果について、定量的な評価ができていないと言いが難い。そこで、本研究ではプログラミングの講義において作問課題を実施し、その学習効果を分析するために学習者の単元理解度を測る。さらに、作問課題のためのルーブリックを作成することにより明確化された評価基準に基づいて、作成された問題文について評価して学習者の個別の理解度を測ることができるようになる。最終的に、講義中に取り組んだ練習問題や小テストの平均所要時間と平均正答率を、作問課題の評価点と比較して分析することにより、プログラミング教育における作問課題の実施効果を明らかにする。

3. 作問学習の実証実験

本研究が実証実験をするプログラミングの講義では、学習者にプログラムを作成させる課題を、学習管理システムMoodle³の電子掲示板にこれまでも提出させてきた。今回実施する作問課題では、プログラムを作らせるのではなく、プログラムを作らせる問題文を学習者自身に考えさせる課題を単元ごとに取り組み、これをMoodle上に提出させる。提出された問題文は本研究のルーブリックに基づいて評価することになる。なお、実験の対象者は、プログラミングを履修している、今回の実証実験の対象範囲となる小テストや練習問題に、少なくとも一回以上取り組んでいる学習者29名である。

3.1 プログラミング教育における作問課題のルーブリックの作成

これまで提案されてきたプログラミング教育のためのルーブリックは、プログラミング的思考を評価するためのものだけであり[13][14]、プログラミング作問課題用のルーブリックは未だ示されていない。そこで、本研究ではプログラミングにおける作問課題を評価するためのルーブリックを作成した。このルーブリックは学習者が作成した問題文の質に関して点数化できる必要がある。そこで、問題文の質に関係する評価尺度を4つ作成し、それぞれの評価尺度の達成度を三段階のレベルに定める。このように作成したルーブリックを表1に示す。

表 1 プログラミング作問課題ルーブリック

| 評価の観点 | 評価の尺度 | | |
|-------------------------|-----------------------------------------|----------------------------------|------------------------------------|
| | 3 (十分に満たしている) | 2 (一部は満たしている) | 1 (満たしていない) |
| 作成された問題は単元の理解を深めるものか | 2の尺度を満たし、かつ単元理解の狙いに沿った追加の要素を含んだ問題文である | 作問課題の条件を満たした問題文である | 作問課題の条件を満たしていない問題文である |
| オリジナリティ・創意工夫はあるか | 2の尺度を満たし、かつ独自の工夫点がある | 講義で取り組んだ練習問題とは異なった問題文である | 講義で取り組んだ練習問題と同じ、もしくは少し変えただけの問題文である |
| 問題文の表現として必要な情報提供がされているか | 2と3の尺度のいずれでもない | 解答者のプログラム作成が容易になり過ぎる過剰な情報を提示している | 解答者がプログラムを作成するのに必要な情報が不足している |
| 作成した問題に対しての解答例は適切か | 2の尺度を満たし、かつ作成された問題から作られる最も無駄がないプログラムである | 作問課題の条件を満たした解答例である | 2の尺度を満たさない、もしくはプログラムとして間違っている |

このルーブリックを活用した評価例として、実際に作問課題に取り組ませた際の課題内容と、学習者により作成された問題文(高評価と低評価)を以下に示す。

以下の点の理解を深めるための文章題を作ります。

- ・ 三分岐以上のif文の構造について理解させる
- ただし、以下のすべての条件を満たすプログラムを作らせるような文章題となっていることが条件です。
- ・ 条件分岐ifを使わせること
- ・ if文の中で入れ子構造となるような、三分岐以上の条件分岐を行わせること
- ・ prompt関数を用いて変数へのデータの入力を行わせること
- ・ 結果表示の際に、連結演算子を使用させること

図 1 作問課題の出題例

あなたは友達とクリスマスパーティーをするため、友達みんなで予算4000円で100円ショップにお菓子や飾りなどを買いに来ました。飲食料品、飲食料品以外の物の買った個数を入力すると、合計金額に応じてメッセージが表示されるプログラムを作ってください。(消費税は軽減税率を適用する)

なお、変数名は各自で決め、条件分岐ifを使い、if文の中で入れ子構造になるようにし、メッセージは連結演算子を使って表示するようにしてください。

- ・ 合計金額が4000円以上の場合「〇〇円は買いすぎだね。次は予算内で収めよう！」
- ・ 合計金額が3000円以上4000円未満の場合「〇〇円で予算内だね。いいパーティーができそう！」
- ・ 合計金額が2000円以上3000円以内の場合「〇〇円ならもう少し買えそうだね。買い足そう！」
- ・ 合計金額が2000円未満の場合「〇〇円では少なすぎるかな。もっと買ってよいパーティーにしよう！」

図 2 ルーブリックで高評価となる作問例

バスケの試合での得点を入力し、その得点に応じたメッセージが表示されるプログラムを作りましょう。

10点未満:「もっと頑張ろう」

10点以上20点未満:「その調子」

20点以上:「すごいね」

とポップアップに表示されるようにしましょう。

図 3 ルーブリックで低評価となる作問例

³<https://moodle.org/>

図2の作問例は、ルーブリックに従うと評価点は満点の12となる。指定された入力処理に関する要素だけでなく、算術演算を要件として加えるなど多くのアイデアが盛り込まれていることが高評価につながっている。一方、図3の作問例は6点の評価となる。この問題文は結果表示の際に連結演算子を用いるという条件を満たしていない。加えて、講義中に取り組んだ練習問題の数値を変えただけの内容となっており、独自の工夫が見られないため低評価となっている。

このように、ルーブリックに基づいて、学習者が作成した問題を評価した結果、問題文が工夫されていて、作問課題の条件を満たしているものを高評価、講義で実施した練習問題と同じで独自性が見られないものを低評価とすることができた。このことから、今回作成したルーブリックは作問課題を適切に評価できているといえる。

3.2 作問課題実施による学習効果の検証

プログラミング教育において作問課題を実施した際の学習効果測定のために、小テストと講義内で取り組む練習問題を利用する。はじめに、講義内で取り組む小テストは学習者に課した作問課題の対象単元から作成した。学習効果の評価尺度としては、小テストおよび練習問題における解答の正誤と、正解を導き出すまでの所要時間の二つを用いることにする。なお、解答の正誤は部分点を用いずに正答と誤答という二段階評価とした。また、所要時間については、時間が短ければ短いほど単元に対する理解度が高く、反対に、所要時間が長い場合には単元理解が不十分であるとみなす。その理由として、単元への十分な理解が、最小限の試行錯誤のみで効率的にプログラムを作成することを可能にすると考えられるためである。さらに、この二つの効果尺度に対し、作成したルーブリックに基づく作問課題の評価点との関係を明らかにすることを目指す。

4. 実験結果と考察

4.1 学習者へのフィードバックと作問課題評価点の関係

はじめに、学習者が作成した問題文をルーブリックによって評価し、これを学習者へフィードバックすることが、その後の評価点の向上につながるかを分析した。表2に示したように、作問課題の取り組みを重ねても、平均評価点は向上しなかった。このことから、評価基準としてルーブリックを示し、それに基づいた教員からの採点結果を学習者にフィードバックしても、高評価の問題文を作成するようにはならないことが明らかになった。

表 2 作問課題評価点のクラス平均点の推移

| 実施回 | 1回目 | 2回目 | 3回目 | 4回目 | 5回目 | 6回目 |
|-------|-------|------|-------|------|------|------|
| 平均評価点 | 10.15 | 9.89 | 10.09 | 8.87 | 9.53 | 9.92 |

4.2 作問課題実施による平均所要時間と平均正答率の変化

次に、作問課題の取り組みの有無によって、講義時間内に実施する小テストの解答所要時間のクラス平均値と、正答率のクラス平均値がどのように変化するかを明らかにする。作問課題の実施手順としては対照実験となるように、学習者を「作問課題に取り組む人」と「作問課題に取り組まない人」の二群に分け、作問課題を4回出題した。なお、二群の分け方の影響を排除するため、作問課題を2回実施することに群の分け方を変更した。表3で示したように、1, 2, 4回目の実施回では作問取組者は作問非取組者よりも平均所要時間が短くなり、平均正答率は高くなった。以上のことから、プログラミング教育における作問課題の実施は、学習者の理解度向上に有用であると結論付けられる。

ただし、3回目の実施回のみ作問実施者のほうが長い所要時間、低い正答率という、理解度が悪くなってしまったという結果になった。その原因として、3回目の実施回に関して、筆者が作成した作問課題の出題範囲と小テストの出題範囲のずれが生じており、適正な検証とならなかった可能性が考えられる。

表 3 作問課題実施の有無が理解度に及ぼす影響

| 実施回 | 解答所要時間のクラス平均値(単位:秒) | | 正答率のクラス平均値 | |
|-----|---------------------|--------|------------|--------|
| | 作問取組群 | 作問非取組群 | 作問取組群 | 作問非取組群 |
| 1回目 | 681 | 688 | 0.667 | 0.643 |
| 2回目 | 550 | 718 | 0.933 | 0.583 |
| 3回目 | 774 | 752 | 0.267 | 0.308 |
| 4回目 | 746 | 816 | 0.417 | 0.313 |

4.3 作問課題の評価点が平均偏差値・平均正答率に及ぼす影響

続いて、作問課題の評価点の大小が小テストや練習問題の理解度に及ぼす影響を明らかにする。まず、4回分の作問課題に対する各学習者の平均評価点に基づいて、クラス平均を境界として学習者を上位と下位の二群に分けた。この二群の学習者ごとに、4回分の小テストと練習問題の平均正答率と平均偏差値を算出する。なお、平均所要時間ではなく平均偏差値を使う理由は、小テストや練習問題の難易度が実施回ごとに異なるため、標準化されていない平均値では単純比較できないからである。そこで、理解度に相当する所要時間に対して学習者の偏差値を実施回ごとに計算し、すべての実施回での平均値を求めた。

表4に示した結果から、平均評価点の上位群に属する学習者は、下位群の学習者よりも平均正答率、

平均偏差値共に高くなるという傾向がみられた。また、作問課題の平均評価点と平均正答率の相関係数は+0.66、平均評価点と平均偏差値の相関係数は0.68となり、作問課題の評価点と小テストおよび練習問題の理解度との間に相関があることが認められた。このことから、作問課題の評価が高くなる人、つまり、作問課題がよくできている人ほど単元への理解度が高くなるといえる。

表 4 作問課題評価点の上位群・下位群の平均偏差値と平均正答率

| 評価点(上位群) | 平均正答率 | 平均偏差値 | 評価点(下位群) | 平均正答率 | 平均偏差値 |
|----------|-------|-------|----------|-------|-------|
| 11.50 | 1.00 | 66.38 | 9.75 | 0.50 | 43.17 |
| 11.50 | 1.00 | 55.63 | 9.75 | 0.50 | 43.02 |
| 11.25 | 0.75 | 54.73 | 9.75 | 0.00 | 35.65 |
| 11.25 | 1.00 | 54.82 | 9.75 | 0.50 | 50.68 |
| 11.25 | 0.75 | 52.30 | 9.67 | 0.50 | 45.94 |
| 11.00 | 0.75 | 52.74 | 9.50 | 0.75 | 52.56 |
| 10.50 | 0.75 | 54.24 | 9.50 | 0.75 | 47.94 |
| 10.50 | 1.00 | 59.73 | 9.25 | 0.25 | 40.62 |
| 10.50 | 0.75 | 52.00 | 9.25 | 0.50 | 45.45 |
| 10.33 | 0.67 | 46.29 | 9.25 | 0.75 | 51.58 |
| 10.00 | 1.00 | 55.78 | 8.75 | 0.50 | 48.95 |
| 10.00 | 0.67 | 53.77 | 8.25 | 0.50 | 44.22 |
| 10.00 | 0.75 | 51.47 | 6.75 | 0.50 | 47.43 |
| 10.00 | 0.75 | 58.32 | 6.33 | 0.00 | 27.74 |
| 10.00 | 1.00 | 56.10 | | | |
| 上位群平均 | 0.84 | 54.96 | 下位群平均 | 0.46 | 44.64 |

5. 研究のまとめ

本研究では、作問課題の実施がプログラミング教育における単元の理解度向上に関してどのような効果をもたらすか、ということについてルーブリックを用いた評価基準の明確化と、講義時間中に取り組む小テストと練習問題の解答にかかる平均所要時間と平均正答率に基づいて明らかにした。

その結果、学習者を作問課題の取り組みの有無から二群に分けると、作問取組群のほうが作問非取組群よりも平均所要時間が短く、平均正答率は高くなり、単元に対する理解度が高くなることが確認された。また、作問課題への取り組みとして、評価点が高い学習者は、小テストや講義時間内に取り組む練習問題の解答所要時間が短くなり、正答率が高くなるという関係性が相関係数によって確認できた。このことから、作問課題の評価点が高い学習者の理解度に与える影響について定量的に明らかにすることができた。

今回の実証実験では、プログラミングの全15回の講義のうち、作問学習を行うのに適した6つの単元において作問学習を行った。しかし、学習効果を図る際に、この実施回数では十分とはいえない。そのため、作問学習の実施回数を増やすことでより多くのデータを収集し、より信頼度の高い学習効果の検証を行うことが今後の課題となる。

参考文献

- [1] 利根川裕太, 佐藤智, 小学校プログラミング教育がよくわかる本, 翔泳社, 2017.
- [2] 文部科学省, 小学校プログラミング教育に関する資料, https://www.mext.go.jp/a_menu/shotou/zyouhou/detail/1416328.htm, (参照 2019-12-23).
- [3] 文部科学省, 小学校段階におけるプログラミング教育の在り方について(議論の取りまとめ), https://www.mext.go.jp/b_menu/shingi/chousa/shotou/122/attach/1372525.htm, (参照 2019-12-20).
- [4] 文部科学省, 教育課程部会 幼児教育部会(第2回)配付資料, https://www.mext.go.jp/b_menu/shingi/chukyo/chuko3/057/siryu/1364728.htm, (参照 2020-2-5).
- [5] 竹中真希子ら, “作問学習を取り扱った先行研究に関する基礎的研究”, 大分大学教育学部研究紀要 第40巻1号, pp.132-148, 2018.
- [6] 平嶋宗, “「問題を作ることによる学習」の分類と知的支援の方法”, 情報システム情報学会研究報告 第20巻3号, pp.3-10, 2005.
- [7] 中野洋二郎, 滝井章, 坪田耕三 編著, 子どもが問題をつくる—小学校算数教科問題づくりの授業, 東洋館出版社, 1999.
- [8] 山岸芳夫, “「プログラミング基礎」における作問学習の実践”, 工業教育研究 26号, pp.91-100, 2017.
- [9] 新開純子ら, “プログラミング教育における作問活動の試み”, 教育システム情報学会, 第39回全国大会, 講演論文集, pp.153-154, 2013.
- [10] 笹栗俊之, 小・中・高等学校でのプログラミング教育実践, 九州大学出版会, 2019.
- [11] 稲垣忠 編著, 教育の方法と技術, 北大路書房, 2019.
- [12] 平嶋宗, “作問学習に対する知的支援の試みと実践”, 科学教育研究 43巻2号, pp.61-73, 2019.
- [13] Benesse, 「プログラミングで育成する資質・能力の評価規準(試行版)」徹底解説, <https://beneprog.com/2018/07/13/standard/>, (参照 2020-2-5).
- [14] 西野和典ら, “プログラミング的思考を評価するルーブリックの作成とその試用”, 教育システム情報学会 第43回全国大会, pp.141-142, 2018.