The role of assessment and reward in non-formal computing settings (Work in progress)

Oliver Quinlan
oliver@raspberrypi.org
Raspberry Pi Foundation
Cambridge, UK

Sue Sentance
sue@raspberrypi.org
Raspberry Pi Foundation
Cambridge, UK

Jonathan Dickins
jonathan.dickins@raspberrypi.org
Raspberry Pi Foundation
Cambridge, UK

Rik Cross
rik@raspberrypi.org
Raspberry Pi Foundation
Cambridge, UK

ABSTRACT

Non-formal learning settings enable self-selected groups of children to participate in out-of-school activities related to their interests, with participation primarily governed by intrinsic motivation. In contrast, formal learning incorporates external rewards through assessment activities which engender extrinsic motivation. Here we describe an exploratory study carried out in computing clubs for primary-aged children, held outside the formal curriculum in schools and libraries. The objectives of the study, involving 12 clubs and 115 children aged 6 to 11, were to investigate the feasibility of introducing assessment opportunities to computing clubs and to consider the impact of rewards in this context. Children worked on programming projects, and were given quizzes to complete at the end of club sessions, with children at some of the clubs being given both quizzes and 'rewards' charts, with badges for completing the projects. We found that quizzes with badges were favourably received by both children and educators. Results indicated that including badges as a form of tracking and rewarding progress significantly increased enjoyment of the assessment activity. This study is part of a larger study looking at learning outcomes in computing clubs, and has implications for others in non-formal computing settings.

CCS CONCEPTS

• **Social and professional topics → Informal education**; *Student assessment*; *K-12 education.*

KEYWORDS

assessment,computing clubs, motivation, non-formal learning

ACM Reference Format:

Oliver Quinlan, Sue Sentance, Jonathan Dickins, and Rik Cross. 2019. The role of assessment and reward in non-formal computing settings (Work in progress). In *14th Workshop in Primary and Secondary Computing Education (WiPSCE'19), October 23–25, 2019, Glasgow, Scotland Uk*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3361721.3362111

WiPSCE'19, October 23–25, 2019, Glasgow, Scotland Uk

1 INTRODUCTION

There is a growing interest in computing in the curriculum worldwide, but for many countries the main vehicle to facilitate young people learning computing is an out-of-school club or holiday activity. Non-formal learning settings such as clubs vary widely and are often run by volunteers. Generally non-formal settings do not include formal assessment activities. However, in order to measure the effectiveness of computing clubs, and their value in the computing education eco-system, it is interesting to know what learning outcomes are being achieved by children through attendance. In addition, maintaining children's interest in attending a club may be impacted by adding extrinsically motivating elements to the intrinsic motivation we assume children have in order to attend.

Part of a larger study around the effectiveness and impact of non-formal computing settings, this paper focuses on a pilot of assessment and reward activities in non-formal learning. We carried out a study in 12 clubs in the United Kingdom in 2018, with the following research question: *Does the use of rewards in addition to assessment activities in non-formal computing settings increase engagement in the activities?*

We first consider related work around non-formal computing, motivation and assessment, then describe the study and findings. Finally we look at the implications of the research for practitioners running computing clubs.

2 BACKGROUND

2.1 Non-formal learning

With computing becoming part of the curriculum in many countries, there is little research in computing education that specifically looks at the difference between formal and non-formal learning contexts. However, a considerable amount of computing teaching is taking place in computing clubs and other non-formal contexts. Some researchers identify a difference between *non-formal learning* and *informal learning*, whereby the latter can be described as opportunistic, unplanned [1] and learner-centric. In contrast non-formal learning can be defined as taking place *"in a planned but highly adaptable manner in institutions,organizations and situations beyond the spheres of formal or informal education"* [3, p.173]. Thus, non-formal learning is intentional, but outside the classroom, and educators may be untrained, often volunteers.

2.2 Motivation

Motivational theories attempt to understand what energises learners towards which activities or tasks [9]. When we refer to someone as being motivated, we mean that the person is trying hard to accomplish a certain task. Intrinsic motivation refers to the individual's participation in activities for personal contentment, enjoyment, curiosity and satisfaction that stem directly from the act of participation without any external reward anticipation [11].

In contrast, extrinsic motivation refers to behaviours and actions initiated by reasons other than an individual's contentment and usually refers to activities such as rewards, benefits and gains or punishments.

2.3 Assessment in computing

In K-12 education a range of common assessment methods are used for formative assessment in computing including self and peer assessment, automated tools, parsons puzzles, rubrics, concept maps and multiple-choice questions or quizzes[12]. Multiple-choice questions are commonly used for a variety of reasons: objectivity [7], number of students [7] and provision of feedback [10].

Although some research points to the fact that multiple-choice questions and quizzes can be used to test low-level skills [14], this does not need to the case, as questions can be constructed to assess higher levels of knowledge on the Bloom's taxonomy [15].

Quizzes have been shown to increase learner engagement, particularly when timed to be given directly after the learning [5].

3 THE STUDY

The intention of the study was to establish whether having rewards (stickers or badges) as well as quizzes was effective in this setting.

3.1 Context

The Code Club network of computing clubs was established in 2012 with 13,000 clubs worldwide. The clubs are held in schools, libraries and community venues across the world, and are run by volunteers and staff at the venues where they take place. Clubs usually operate on a weekly basis for around an hour each session.

3.2 Research Design

For this study, participating clubs were asked to use a series of six projects designed to teach programming skills to beginners using the Scratch programming language. The clubs were divided into two groups, with one using the quizzes only and one using the same quizzes and a rewards chart that children completed with stickers (badges) linked to completing each project.

3.3 The intervention materials

3.3.1 The quizzes. A series of multiple choice quizzes were designed to test the children's comprehension of the learning objectives for each project. Figure 1 illustrates how these were presented to the children.

Each quiz had three multiple choice questions, with the final quiz containing a further three questions, designed to assess computational thinking skills across the six projects.

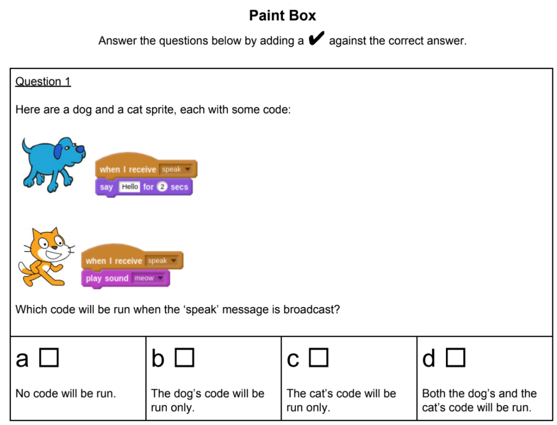

Figure 1: Example quiz question

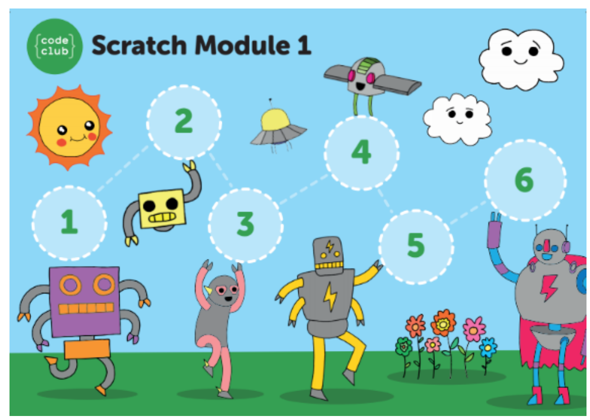

Figure 2: Reward chart for quiz completion

Questions were of different types, for example, some requiring the child to predict what a piece of code would do [13], others involving tracing of a snippet of code [8], and others where children needed to specify which block of code could be used to reach a particular outcome.

3.3.2 Reward chart and badges Âl' Raspberry Pi Foundation. The rewards chart (shown in Figure 2) was given to half the schools in the survey. Children collected a badge (sticker) for each completed project. Earning badges was not related to performance on the quizzes.

3.4 Data collection

Data was collected in a variety of ways:

- **Project log**: Clubs were provided with a project log for each child to complete.
- **Survey**: A survey was used to collect data from children at the end of the final session. It included questions around engagement, difficulty and utility, with all questions using a 5 point Likert scale. Adults were asked to set aside time in the final session for the children to complete it.

- **Interviews**: Qualitative data was collected through visits to four of the clubs where sessions were observed and club leaders interviewed, with telephone interviews with the remaining clubs. Semi-structured interviews were structured around three topics: the implementation of the study, the adults' views on the project, and children's response to the intervention.
- **Performance data**: Quantitative data was collected on the performance of children on the quizzes.

3.5 Participants

The study took place in clubs in the autumn term of 2018. 12 clubs participated in the trial, all based in the UK. Nine of the clubs took place in schools and three in libraries, with five run by staff of the venues and seven by volunteers. The children taking part were all aged between 6 and 11, with some clubs catering for a particular year group and others being open to children of a wider range of ages. Data was returned for 115 children across the project, 59 boys, 37 girls and 19 who chose not to disclose their gender. The number of children at each club ranged from 3 to 16, with an average of 9.3.

3.6 Recruitment and sampling

Clubs were recruited via the annual survey, in that a question was added eliciting participation.

The clubs were randomly allocated to the two groups: quizzes only, or the quizzes and the badges versions of the project. Two clubs dropped out early in the trial, one reporting a lack of enthusiasm from the children to take part and one due to some organisational issues with the club.

This resulted in five clubs using the quizzes only and seven using the quizzes and badges.

3.7 Ethics

This project was planned with reference to the British Education Research Association (BERA) ethical guidelines for educational research [2]. No personally identifying information was collected about children. Parental consent was sought by each club for participating children.

3.8 Data analysis

3.8.1 Survey data. Survey results were compared across the different clubs. For each group data were analysed to identify the differences in the two groups on the following variables: enjoyment (measuring interest), and utility (whether children felt quizzes helped with their learning). Statistical significance for comparisons was tested using a Wilcoxon Rank Sum Test.

Children were asked three questions relating to whether they liked the quizzes, whether they liked earning the badges (if in a group that used them), and whether the quizzes helped them to learn. This was collected as agreement rated on a five-point Likert scale. Responses to the three questions were averaged to give an overall feedback rating.

Children were also asked about the extent to which the quizzes had developed their skills in five areas relating to success in programming. These responses were also averaged to give a learning rating.

3.8.2 Interview data. Detailed notes were taken on interview recordings under the headings of the structured questions that were asked. The resultant data was analysed using thematic analysis [6], creating a series of codes. The initial analysis was deductive, using themes derived from the interview questions:

- Administration of the intervention
- Attitudes towards the quizzes
- Attitudes towards the badges
- Use of the provided answers (such as marking the quizzes or going through answers with the children)
- Adults' use of the quizzes (such as for understanding learning)

A further analysis was conducted to explore themes that emerged inductively from the data [4], giving two additional themes:

- Challenges with time taken to complete quizzes
- Discussion or reflection activities after completing quizzes

Interview notes were iteratively revisited and coded against these seven themes. These themes were combined with the deductive themes for reporting. In the next section we detail the results of the project.

4 RESULTS

4.1 Survey findings

4.1.1 Enjoyment. On the question of whether children liked doing the quizzes a rating of a 4 or 5 was classified as a positive response to the activity. 62% of the children rated that they liked doing the quizzes. There was a statistically significant difference between the proportion of children in each group ('quizzes' and 'quizzes and badges') who liked doing the quizzes ($W=998$, $p<0.1$). 51% of children who only completed the quizzes reported liking them, whereas 72% of those who also received badges did so.

There was a significant difference between the groups on the feedback metric ($W=1094$, $p<.01$), with the average score given by those who received badges 0.9 higher (on a scale of 1–5) than those who did not. The proportion of children who expressed that they liked the quizzes ranged from 27% to 100% , a range of 73 percentage points with a standard deviation of 0.23.

4.1.2 Perceived utility. We asked the children whether they thought that the quizzes had helped them to see what they had learned. 55% of the children rated this question as either a 4 or a 5, which we interpreted as agreement. There was a difference of 22 percentage points between the groups, with 67% of those who received the badges as well as using the quizzes agreeing, and 45% of those who only used the quizzes agreeing.

4.2 Results from interviews

As detailed in Section 3.8.2, we analysed data around seven distinct themes. Here we discuss only the themes related to attitudes and to learning, as these gave the most interesting insights related to our research question.

4.2.1 Attitudes towards the quizzes. Many adults fed back that the length, format, and complexity of the quizzes was appropriate in their informal club environment. Only one adult reported that the quizzes were too formal for the ethos of their club, but they took copies of them to use in their computing lessons.

Most adults reported that the level of challenge of the quiz questions was appropriate. Three adults thought the quizzes were too hard for the children in their club. In two cases they expressed that the fact that the quizzes presented concepts in a different context to the Scratch projects made them difficult for the children.

Adults mostly reported that children were happy to complete the quizzes. Some particularly noted children's enthusiasm. Those that noted they were too hard also said that the children did not enjoy them.

Six of the 12 adults reported that they felt the quizzes encouraged the children to reflect on what they had learned by completing the projects. Others said they did not, with two reporting that they did not feel the children engaged with the quizzes strongly enough for this to happen. One adult fed back that the projects clubs use are very focused on achieving an end product. Although they felt this to be motivating, they expressed that an opportunity for reflection on learning before moving the focus to the next product to be built was beneficial to the children, and even helped them to transfer skills learned during one project to the next.

4.2.2 Attitudes towards the badges. Six of the seven adults whose clubs used the badges were positive about this aspect of the project. One expressed that they felt the badges were more motivating and appropriate for younger children. One adult said that they had already been giving out stickers, taking a similar approach.

4.2.3 Adult use of quizzes. Four adults said that they made use of the quizzes to understand what the children had learned from the projects, and the areas they found problematic. Other adults said they did not use the quizzes for this, and rather saw them as an exercise for the children. One adult reported that they could tell how well the children had learned concepts by their engagement while programming, whether they had problems, and the questions they asked. Adults tended to describe the utility of the quizzes as a formative assessment tool, rather than a summative one.

4.2.4 Discussion or reflection activities. An emerging theme that was present in four clubs related to the use of quizzes as opportunities for discussion or reflection activities. Activities undertaken ranged from discussing the answers with the children to discussing questions the children had found challenging. One adult expressed the view that the discussions resulting from the quizzes were the most beneficial part of the project.

5 CONCLUSION AND FUTURE WORK

In this paper we set out to investigate how assessment activities work in computing clubs. We have found the use of assessment alongside a reward system engages young people more than assessment alone in these contexts. We have found a range of different intervention approaches by the educators and volunteers in our study, reflecting the diversity of non-formal learning settings. Despite this, there was a positive impact of the use of both quizzes and rewards, particularly where the participant adults engaged the children meaningfully with the activity. We will use our findings to inform future work on a larger scale.

With an emerging but patchy roll-out of computing in the curriculum around the world, the role of non-formal learning cannot be ignored. In countries where computing is mandatory in primary and lower secondary education it can complement and extend formal education. Elsewhere it may be the only opportunity available for children to learn computing skills (including programming) before they reach more advanced education. Therefore we need to thoroughly investigate the format of non-formal learning to find out what is most effective in engaging and educating children. There is a substantial research agenda in this area beyond that that we have started to address in this paper, including what pedagogical strategies work well in non-formal learning settings, how volunteers and untrained adults can support children effectively to learn computing, and how we can impact diversity and inclusion in self-selecting non-formal settings.

REFERENCES

[1] Andrew Begel and Andrew J. Ko. 2019. Learning Outside the Classroom. In *The Cambridge Handbook of Computing Education Research*, Sally A Fincher and Anthony V Robins (Eds.). Cambridge University Press.

[2] British Educational Research Association [BERA]. 2018. Ethical Guidelines for Educational Research, fourth edition. (2018). https://www.bera.ac.uk/researchers-resources/publications/ethicalguidelines-for-educational-research-2018

[3] Haim Eshach. 2007. Bridging in-school and out-of-school learning: Formal, non-formal, and informal education. *Journal of science education and technology* 16, 2 (2007), 171–190.

[4] B. Glaser and A. Strauss. 1967. *The discovery of grounded theory: Strategies for qualitative research.* Aldine Publishing Company.

[5] Alice F Healy, Matt Jones, Lakshmi A Lalchandani, and Lindsay Anderson Tack. 2017. Timing of quizzes during learning: Effects on motivation and retention. *Journal of Experimental Psychology: Applied* 23, 2 (2017), 128.

[6] Udo Kuckartz. 2014. *Qualitative text analysis: A guide to methods, practice and using software.* Sage. https://doi.org/10.4135/9781446288719

[7] William L Kuechler and Mark G Simkin. 2003. How well do multiple choice tests evaluate student understanding in computer programming classes? *Journal of Information Systems Education* 14, 4 (2003), 389–400.

[8] Raymond Lister, Elizabeth S. Adams, Sue Fitzgerald, William Fone, John Hamer, Morten Lindholm, Robert McCartney, Jan Erik Moström, Kate Sanders, Otto Seppälä, Beth Simon, and Lynda Thomas. 2004. A Multi-national Study of Reading and Tracing Skills in Novice Programmers. In *Working Group Reports from ITiCSE on Innovation and Technology in Computer Science Education (ITiCSE-WGR '04).* ACM, New York, NY, USA, 119–150. https://doi.org/10.1145/1044550.1041673

[9] Paul R Pintrich. 2003. A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of educational Psychology* 95, 4 (2003), 667.

[10] Tim S. Roberts. 2006. The Use of Multiple Choice Tests for Formative and Summative Assessment. In *Proceedings of the 8th Australasian Conference on Computing Education - Volume 52 (ACE '06).* Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 175–180. http://dl.acm.org/citation.cfm?id=1151869.1151892

[11] Richard M. Ryan and Edward L. Deci. 2000. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology* 25, 1 (Jan. 2000), 54–67. https://doi.org/10.1006/ceps.1999.1020 14717.

[12] Sue Sentance, Cynthia Selby, and Maria Kallia. 2018. Assessment in the Computing classroom. In *Computer Science Education: Perspectives on Teaching and Learning in School*, Sue Sentance, Erik Barendsen, and Carsten Schulte (Eds.). Bloomsbury Academic, London, 151–166.

[13] Sue Sentance, Jane Waite, and Maria Kallia. 2019. Teachers' Experiences of using PRIMM to Teach Programming in School. In *Proceeding of the 50th ACM Technical Symposium on Computer Science Education (SIGCSE '19).* ACM, New York, NY, USA, 561–566. https://doi.org/10.1145/3287324.3287477 00023.

[14] Shuhaida Shuhidan, Margaret Hamilton, and Daryl D'Souza. 2010. Instructor perspectives of multiple-choice questions in summative assessment for novice programmers. *Computer Science Education* 20, 3 (2010), 229–259. https://doi.org/10.1080/08993408.2010.509097 00030.

[15] Karyn Woodford and Peter Bancroft. 2005. Multiple choice questions not considered harmful. In *Proceedings of the 7th Australasian conference on Computing education-Volume 42.* Australian Computer Society, Inc., 109–116.