International Journal of Child-Computer Interaction 2 (2014) 85–92

Contents lists available at ScienceDirect

International Journal of Child-Computer Interaction

journal homepage: www.elsevier.com/locate/ijcci

From pen-and-paper content to educational math game content for children: A transfer with added difficulty

Marie Maertens [a,b,*], Mieke Vandewaetere [a,b], Frederik Cornillie [b], Piet Desmet [b,c]

[a] CIP&T, Center for Instructional Psychology and Technology, KU Leuven, Dekenstraat 2, 3000 Leuven, Belgium

[b] ITEC – iMinds – KU Leuven - Kulak, Interactive Technologies, KU Leuven – Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium

[c] Franitalco, Research on French, Italian and Comparative Linguistics, KU Leuven, Blijde-Inkomststraat 21, 3000 Leuven, Belgium

ARTICLE INFO

Keywords:
Educational games
Learning effectiveness
Item calibration
Format effect
Methodology

ABSTRACT

The use of educational games is nowadays well-known, although its effectiveness in terms of learning and motivation has not been firmly corroborated yet. In this research, the focus is on the use of adaptive item sequencing in an educational math game (Monkey Tales) as a way to provide learning content that is adapted to the knowledge level of the learner or player. One way to provide this is adjusting the difficulty level of the content that is presented hence allowing for optimal learning during gameplay. In order to realize this, accurate estimates of the learners' skills are necessary, as well as accurate estimates of variables in the gaming environment that determine how much difficulty learners may experience while completing the items. The goal of this study is to compare difficulty ratings from users and experts in order to acquire reliable estimates of the difficulty of the math rules offered in Monkey Tales. The results suggest a double format effect: not only the gaming format adds difficulty to educational content, but also the format by which educational content is presented in a game is likely to add difficulty which affects learners' in-game performance.

1. Introduction

Although the use of games for educational purposes is far from new, the latest years have sparked a renewed interest in the phenomenon. In stark contrast with this, however, empirical research over the years has yielded surprisingly little insight about its effectiveness in terms of learning and motivation [1–3]. A possible explanation for the lack of clear findings is the infrequent interaction between the disciplines of, on the one hand, instructional design research and, on the other hand, game design and development. Consequently, sound instructional design principles and concepts (like adaptive instruction) are typically not incorporated in educational games. In this paper, we focus on the use of adaptive item sequencing or adaptive selection of items as a way to provide learning content that is adapted to the skills of the player (i.e. the learner; [4]). One basic example of adaptive item sequencing is that, if a learner fails to complete a task, the subsequent task is more easy and when a learner successfully completes a task, the subsequent task is more difficult, hence increasing the challenge.

However, in order to realize more fine-grained adaptivity, we require, on the one hand, accurate measures of the skills and knowledge of learners, and, on the other hand, a comprehensive account of how educational game environments may determine how much difficulty learners experience while completing items [5]. The latter account needs to include two types of estimates. First, there are the estimates of the inherent difficulty of the learning content, and second, there are the estimates of the 'additional' difficulty that may be imposed by the format in which the learning content is offered. This format comprises, among other things, gameplay elements such as competition with opponents or time pressure. When offered in a game environment, learning content might elicit other mental processes and hence impose higher or lower cognitive load, become more or less difficult and are less or more effective when gaming elements come into play. For the design of effective educational games, then, it is crucial that these estimates are obtained, since learners are likely to lose interest or become frustrated when, learning content is too easy, or gameplay is not sufficiently challenging for them, or content and gameplay are too difficult. Having at their disposal both the inherent difficulty of learning materials and the additional difficulty as imposed

* Corresponding author at: ITEC – iMinds – KU Leuven - Kulak, Interactive Technologies, KU Leuven – Kulak, E. Sabbelaan 53, 8500 Kortrijk, Belgium. Tel.: +32 0 56 24 61 18; fax: +32 56 24 60 52.

E-mail address: marie.maertens@kuleuven-kulak.be (M. Maertens).

http://dx.doi.org/10.1016/j.ijcci.2014.04.001
2212-8689/

Math rule 6: Subtraction with 3 full tens to 100 (e.g. 70-20 = …, 31-10 = …)

Please indicate the degree of difficulty for this rule, for learners of the third grade:

0	1	2	3	4	5	6	7	8	9	10
Very **easy**. Almost no learners will make errors.								Very **difficult**, almost all learners will make errors		

Fig. 1. Exemplary math rule with examples between brackets for the difficulty ratings – experts.

by gameplay formats, developers and instructional designers can create adaptive experiences that allow for optimal learning and challenge.

To provide learning content that is based on reliable estimations of difficulty levels, a calibration study is required at the earliest stages of game development. However, at the time of content calibration, game development is often still in the stage of conceptualization and design. Hence, the methods used for content calibration are limited to the use of pen-and-paper (P&P). Consequently, adaptive gameplay is created based on difficulty levels that were obtained outside the game environment, assuming that what works in regular education will also work in a gaming content. However, when offered in a game environment, learning content (like math items) might elicit other mental processes and hence impose higher or lower cognitive load and work less or more effective when gaming elements like opponents, time pressure and competition come into play.

To test the hypothesis that both learning content and gameplay determine the challenge of educational games, we carried out a study in which we calibrated the difficulty of content presented in the educational game for children *Monkey Tales* (www.monkeytalesgames.com). This game was designed for mathematics practice in elementary school and is based on a sequence of math rules. The game requires learners to solve 3D puzzles, each of which comprises a mini-game. Each mini-game features one math rule presented to the learners as a set of math items (e.g. item '$8 \times 6 = \cdots$' for the math rule 'table of six'). The game features six types of mini-games, and learners must perform different and potentially more challenging tasks in each mini-game. For example: in one mini-game, a ball must be thrown at the correct answer, but in another mini-game, learners must navigate to the correct answer while evading slime and shoot at the correct answer. The goal of the current study was to provide appropriate and reliable estimates of the difficulty of the math rules presented in the game, taking into account the following parameters: (1) context in which the rule is rated: by teachers/experts or by learners/users; (2) environment in which the rule is provided: P&P or game environment; (3) type of mini-game in which the rule is provided: whether gameplay features confound the estimates of difficulty; and (4) intended users of the rule: the math ability of learners.

2. Method

The estimation of math rule difficulty was based on four procedures: (1) expert ratings: math experts indicated the difficulty level of the items on a scale from 0 to 10, by providing an estimation of the proportion of students that would be able to correctly solve the item; (2) user ratings on P&P: items per rule were provided to third grade pupils and they were asked to solve the rule and to indicate the experienced difficulty of the items on a scale from 1 to 9; (3) user ratings during gameplay: after completing a set of math items in the game, the same third grade pupils indicated the difficulty they experienced in solving the items; on a scale from 1 to 9. The user and expert ratings are self-reported and must therefore be considered as subjective. This makes the fourth method more objective: (4a) proportion correct responses per rule, during gameplay (i.e. at the end of gameplay, the proportion of learners that had correctly solved the rule was computed) and (4b) proportion correct responses per rule as offered in the P&P test. This is a more objective indication of the difficulty level of the rule: the higher the proportion correct, the more easy the rule is to solve [6].

The content of Monkey Tales is defined by math rules (e.g., 'Table of 6', 'Addition to 100', 'Odd and even', etc.). Each math rule is operationalized by a predefined number of items that can be offered to Monkey Tales users. Items thus do not differ with respect to the underlying math rule, they only differ in the numbers that are used. The items belonging to the same math rule have been categorized in the standards of the Flemish curriculum based on which Monkey Tales was developed. Items that belong to the same level of difficulty (as categorized in the standards) can become more difficult depending on the context (in this case, the mini games) in which they are offered. We can thus assume that items having the same rule as origin will also have the same difficulty level. Therefore it is not necessary to estimate difficulty on the item level as estimations of difficulties at the rule level are sufficient to provide reliable estimates.

2.1. Method expert ratings

Five teachers of the third grade in primary education in Flanders participated in the expert rating study. All teachers were female and had between 4 and 25 years of experience in the third grade. A P&P questionnaire was given to the experts and they could fill in the questionnaire whenever they wanted. They were made aware of the purpose of their ratings and the grade for which they rated the content (third grade). Experts were asked to complete their profile with respect to years of teaching experience in the grade for which she rated; gender; age; type of education; and specification of didactical materials used for teaching (like handbooks). The ratings of the content were organized as follows (see Fig. 1): for the 44 math rules as offered in the Monkey Tales game for third graders, one or two examples were provided. As the experts were asked to indicate the proportion of third grade pupils that would be able to solve the item correctly at the beginning of the third grade, a scale from 0 to 10 was selected in terms of interpretability (i.e. how many children out of 10 would have difficulties).

2.2. Method user ratings and proportion correct per rule

Three schools (Flemish education) agreed to participate resulting in access to seven third grade classes (between 8 and 9 years old). Administration of the study took place in January, hence at the middle of the school year. All 140 learners received an informed consent for their parents and 101 learners (and their parents) consented to participate in the study resulting in a response rate of 72.14%. From the 101 learners, 67 actually played the Monkey Tales game resulting in a final response rate of 47.68%. Overall drop-out