Journal of Information Technology Education: Research

An Official Publication of the Informing Science Institute
InformingScience.org

JITEResearch.org

Volume 17, 2018

REVIEW OF FEEDBACK IN DIGITAL APPLICATIONS – DOES THE FEEDBACK THEY PROVIDE SUPPORT LEARNING?

Betty Tärning | Lund University, Lund, Sweden | Betty.tarning@lucs.lu.se

ABSTRACT

Aim/Purpose The goal of this paper is to examine digital applications used in Swedish schools and whether they fulfill their potential as support for learners. This is done by examining the kinds of feedback they provide and discussing if this feedback supports learning or not.

Background The paper targets one aspect regarding which educational apps can be of high value for learners and teachers, namely the feedback they provide. The paper also addresses the need for supportive feedback and reviews 242 apps with respect to what types of feedback they provide.

Methodology A sample of apps used in primary school was collected via email to schools in Sweden. The author evaluated each app with respect to what kind of feedback it provided. The article concerns both positive and negative feedback, with a focus on negative. The following types of feedback were evaluated; verification feedback, corrective feedback, elaborated feedback, encouraging feedback and result feedback.

Contribution This paper contributes to knowledge regarding how most apps only contain verification feedback (telling the student whether their answer was correct or not). In order to help a student while learning, verification feedback is not enough. Rather, previous research has shown that explanatory feedback is more beneficial for learning.

Findings Seventy-seven percent of all apps contained verification feedback, and only 12 % provided the student with some type of explanation as to why their answer was incorrect. Looking at previous research, this is not desirable if one wants the app to support learning and not only act as a testing device. Fifty-five percent of all apps also contained some type of encouragement, but none of this encouragement addressed the task or the effort the learners put into the task - something that would be preferable from a learning perspective.

Accepted by Editor Keith A. Willoughby │Received: April 12, 2018│ Revised: July 9, July 30, 2018 │ Accepted: August 17, 2018.

Cite as: Tärning, B. (2018). Review of feedback in digital applications – Does the feedback they provide support learning? *Journal of Information Technology Education: Research, 17,* 247-283. https://doi.org/10.28945/4104

Recommendations for Practitioners There is much to be gained for developers of educational software if they would make more use of the feedback in educational apps. As for now, the feedback is primarily suited for testing and not for learning. For users of apps (teachers, parents, and children) this paper shows that feedback can be and is an important factor to evaluate before deciding if the app is "worth" spending time on.

Recommendation for Researchers The research describes different types of feedback and their (dis)advantages.

Impact on Society The paper stresses that most feedback represented in apps today corresponds to a behavioristic approach comparable to instrumental conditioning by means of reinforcement. In essence, most apps miss the opportunity of treating the learner as an active and constructive being who would benefit from more nuanced feedback.

Future Research Previous research has shown that elaborated feedback is more beneficial for learning, but more research needs to be done here, the amount of elaborated feedback will most likely affect varying student groups and varying tasks in different ways. And more importantly, how can we make the students pay attention to and act upon the feedback provided to them.

Keywords digital applications, verification feedback, corrective feedback, elaborated feedback, encouraging feedback, result feedback

INTRODUCTION

Digital applications are today part of everyday school life; the number of educational apps has grown immensely over recent years. However, evaluation of the applications with respect to their effects on learning has generally been neglected.

As a consequence, it is difficult for teachers, parents and students to select which apps to use. In the Swedish context, there are web pages such as *Pappas appar* (http://www.pappasappar.se/) or *Länkskafferiet* (http://lankskafferiet.org/) that describe and score apps. The scoring is done by parents, some of whom are teachers or other school-related staff. The criteria for scoring vary but include whether the app is free to download, whether it has nice graphics, whether one´s own children or class seem happy using it. The guidance is of value, but it is hardly a systematic evaluation of how the applications affect learning.

This paper targets one aspect regarding which educational apps can be of high value for learners and teachers: the feedback provided to the learners.

Feedback is a consequence of performance. The learner receives a response to what s/he does: a response that, in most cases, tells something about the quality of his/her action or answer. For example, learners can be informed simply whether their answer was correct – so-called *verification feedback* – sometimes accompanied by the correct answer. Other types of feedback provide learners with more information: why an answer was correct or incorrect, or a small hint pointing toward the correct answer.

It is well-established that feedback has a large effect on learning (Black & Wiliam, 1998; Hattie & Timperley, 2007; Shute, 2008). Given this, it is striking how relatively little attention feedback has received when it comes to educational apps. Even though a teacher is still unbeatable when it comes to providing individualized feedback, digital systems have a potential that teachers do not. A teacher cannot simultaneously place herself beside every student to provide individualized feedback.

The goal of this paper is to examine the kinds of feedback provided in the educational apps used in Swedish schools today and discuss whether that feedback supports learning.

BACKGROUND

GUIDING THE LEARNER IN EXPLORATION

A debate about the best way to learn has raged for decades. At one end of the scale, we find those who recommend free play, where the learning environment is not structured or designed in any purposeful way (Gray, 2013, cited in Hirsh-Pasek et al., 2015). At the opposite extreme, we find those who believe only in highly structured instruction, where the teacher explains how things work and what the learner needs to know (Hirsh-Pasek et al., 2015). In effect, there are pros and cons to both approaches, but the best solution lies somewhere in between (Schwartz, Tsang, & Blair, 2016).

In *free* or *discovery* learning (Mayer, 2004), the learner explores an environment with little or no guidance. It is up to the learner herself to select, organize and integrate information. An advantage is that the student is free to construct his/her own learning experiences and is forced to take an active role in the learning task. At the same time, free exploration in a complex environment can generate a high cognitive load, detrimental to learning (Sweller, 1994). The problem is especially relevant for novice learners, who lack existing frameworks into which to integrate the new information and who therefore must search the problem space more thoroughly. Kirschner, Sweller, and Clark (2006) argue that free learning makes a poor fit with our cognitive architecture.

The idea that learners should construct their own knowledge is reasonable; however, leaving learners without guidance in that endeavor is not. Many, if not all, learners struggle at one point or another when left on their own (Chi, 2009). A learner may focus on the wrong information from the beginning. Then it becomes hard – if not impossible – to straighten oneself out. The learner needs someone or something to guide him/her in the right direction again. A meta-review from Alfieri, Brooks, Aldrich and Tenenbaum (2011) concludes that direct instruction results in better learning than free play, but that the best learning is achieved through assisted discovery, with the instructor taking a supportive "back seat" role. Mayer (2004) argues that overwhelming evidence should make anyone skeptical of the benefits of pure discovery learning, with experimental evidence all pointing in the direction of having guidance when exploring a learning environment.

Learners need help not to treat new information as something just to memorize and recite. Rote memorization typically does not lead to so-called transfer. The goal should be to train students to be self-regulating learners, taking control of their own learning. Students need to be able to recognize when they understand and when they do not (Bransford, Brown, & Cocking, 2000).

GUIDING THE LEARNER TO EXPLOIT HIS/HER LEARNING POTENTIAL

Guiding and being guided are everyday experiences: we observe how others do things and we act as role models to others – often without knowing. Lev Vygotsky was one of the first to recognize the importance of guidance to learning. That guidance might come from a parent or a more experienced peer. In either case, someone who is more experienced helps someone who is less experienced move from their current *performance* level to their *potential* level: what the individual can do with help. Vygotsky (1980) calls the gap between these the *zone of proximal development*. At first, the learner may need help at every step. Gradually s/he is able to perform some steps independently. Finally, s/he performs the entire activity with no assistance at all. Assisted performance guides the learner toward achieving things s/he could not achieve on his/her own (Gibbons, 2002).

Children are intrinsically motivated to participate in many kinds of activities, but they may not always see why certain activities are important. This is up to adults to explain. Vygotsky observed optimal motivation in children when asked to perform just above their present abilities (their present performance level). This means that a child can be motivated to learn more and make further progress if we provide them with *scaffolding*.

Scaffolding is a broad concept, encompassing all kinds of support provided to a learner in order to back his/her up in his/her learning activities. Feedback is a part of scaffolding – scaffolding provided in response to what the learner does.

Scaffolding

Wood, Bruner, and Ross (1976, p. 90) define scaffolding as:

> [A] process that enables a child or novice to solve a problem, carry out a task or achieve a goal which would be beyond his unassisted efforts. This scaffolding consists essentially of the adult 'controlling' those elements of the task that are initially beyond the learner's capacity, thus permitting him to concentrate upon and complete only those elements that are within his range of competence.

As with any scaffolding, the scaffolding is removed over time, allowing learners to accomplish the same task on their own. Since then, scaffolding has become a well-researched topic, and researchers have discussed which factors or ingredients are important for it (Bransford et al., 2000; Van de Pol, Volman, & Beishuizen, 2010).

Bransford et al. (2000, p. 104) list as factors the following six activities or tasks as possible ingredients in scaffolding:

1. Making sure the learner keeps up interest in the task.
2. Reducing the number of steps needed to solve the task.
3. Motivating and directing the learner to pursue the goal.
4. Pinpointing the differences between the learner's current work and the desired outcome.
5. Reducing frustration and risk.
6. Demonstrating what an ideal performance looks like.

Feedback provided while a learner is working on a task seems to be a key ingredient for successful learning. Lepper, Aspinwall, Mumme, and Chabay (1990) further examined how expert tutors scaffold their learners. They conclude that experts tend to draw the learners' attention to an error, then provide a second chance at the solution – instead of offering corrective feedback. They usually ask the learner questions and avoid explicit directions. Fox (1991) reports a similar pattern.

Scaffolding and feedback intertwine. Scaffolding is the wider concept, including all forms of support given throughout the learning process. Scaffolding can also be a way of preventing a situation before it occurs (i.e. before the learner does something unwanted), and it can also be used to provide targeted support for particular learners or to deliver general instructions to a whole group of learners. On the other hand, feedback is information brought to the learner in response to something s/he has done.

Feedback

In general terms, feedback can be said to be information coming back to a person in response to his/her performance, thoughts or ideas. According to Hattie and Timperley (2007), feedback can provide the learner with corrective information, provide alternative strategies, bring information to clarify ideas, provide encouragement, and provide the learner with correctness regarding their response.

Review studies by Black and Wiliam (1998), Hattie and Timperley (2007) and Shute (2008) show that feedback can help learners to achieve their learning goals better. With that said, feedback *per se* does not ensure good performance. If learners can peek at what is designed as feedback before they have constructed their own answer, there is little effective 'feedback'. The learner could merely copy-&-paste the answer without reflecting at all (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991). Feedback in the form of grades or other markers telling the learner how they are doing in comparison to others is usually not beneficial (Butler, 1987; Kluger & DeNisi, 1996; Wiliam, 2007). In contrast, feedback that contains information about the task and how to do it more effectively supports learning (Hattie & Timperley, 2007).

There are many different forms of feedback. First of all, we have *positive* and *negative* feedback. Positive feedback is information that tells the learner that there is no need for further learning or action; they already made a correct response. Negative feedback tells the learner that there is a discrepancy between his/her performance and his/her learning goal (Schwartz et al., 2016). Unfortunately, negative feedback can have a threatening effect on learners. Learners who are *performance oriented* rather than *mastery oriented* are not likely to see negative feedback as something positive. The feedback tells them that they failed, which they may interpret as an indication that their performance was not good enough and that they are not smart enough. Learners who are mastery oriented are more likely to see the feedback as a chance to improve their learning. To them, negative feedback is meaningful in that it helps them in their goal: to learn and make progress.

This article will be concerned with both positive and negative feedback, but the focus will be on negative feedback.

The amount of information contained in feedback varies from "none" to "too much" (Schwartz et al., 2016). With respect to adequateness, both *amount* and *kind* of feedback vary between groups of learners. Novices generally need more information to correct their answers, compared to more knowledgeable learners.

Kulhavy & Stock (1989) write that good feedback should contain verification (*whether* the answer is right or wrong) and elaboration (*why* the answer is right or wrong). If a learner receives adequate feedback, this can reduce the uncertainty of where s/he stands in relation to the task. Uncertainty often takes attention away from the task itself. Adequate feedback can help reduce cognitive load and provide information useful for correcting misconceptions or inappropriate strategies (Shute, 2008). Good feedback should also be specific and timely: feedback should make clear the difference between learner performance and goal; and it should be delivered in reasonable time, so the learner can correlate the feedback to the task (Schwartz et al., 2016). Feedback should be understandable (Higgins, Hartley & Skelton, 2001; Lea & Street, 1998; Orsmond, Merry & Reiling, 2005; Schwartz et al., 2016), non-threatening (Schwartz et al., 2016), and reasonable (Brockbank & McGill, 1998). Lastly, the learner must be able to see the connection between the feedback and the task; otherwise, s/he will not see the point in using the feedback (Orsmond et al., 2005; Segedy, Kinnebrew, & Biswas, 2013; Wiliam, 2007).

Whether the feedback is adequate for each and every student or presented within a reasonable timeframe are things that will not be discussed here. Of interest, instead, is the amount of information that different types of feedback provide the learner with and whether or not this amount is preferable for learning. The reviewed feedback types can be seen in Figure 1 and will be described and discussed below.

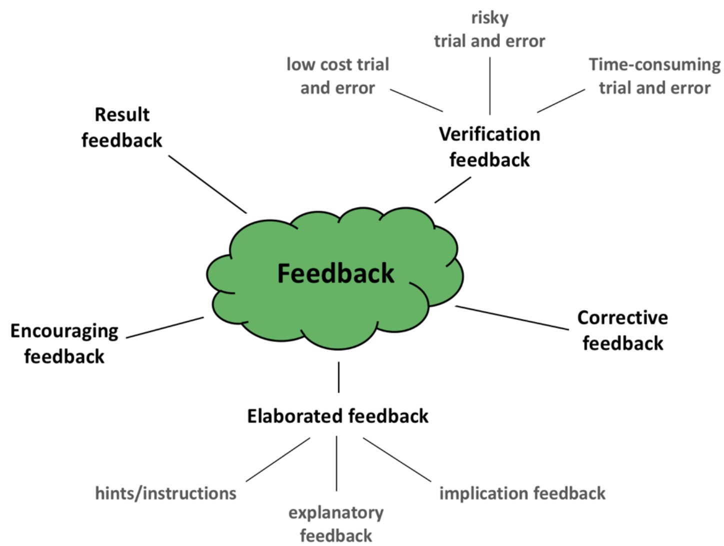

Figure 1. Reviewed feedback types

Verification feedback

A simple form of feedback gives the learner verification of whether his/her answer was correct or incorrect. I will call this *verification feedback* (sometimes also known as *knowledge of result* or *right/wrong feedback*). Verification feedback provides the learner with a sense of knowing whether s/he is on the right track: it can be more or less explicit. Examples are when the learner enters an answer, and the app indicates 'incorrect' or 'correct' via words or symbols, e.g. a red cross vs a green checkmark or a sad vs a happy face (Figure 2). Often such markers are accompanied by a negative or positive sound. A correctly spelled word may be read out loud by the software.

Figure 2. An example of direct verification feedback: the student has answered incorrectly, as shown by the unhappy red faces

Implicit verification looks a little different. Say that a learner solves a crossword and tries to spell a word correctly by dragging a letter to one of the squares. If s/he chooses incorrectly, the letter is automatically removed from the square without explicit sounds or other types of signals indicating that the choice was incorrect. When the learner drags a letter to its right place, the letter stays, indicating that the choice was correct.

Studies on verification feedback are not unanimous concerning its usefulness. Pashler, Cepeda, Wixed, and Rohrer (2005) let their participants learn English translations of Luganda words (e.g., "*leero*" means "*today*"). After an initial training session, participants took a test and received (i) no feedback, (ii) verification feedback, or (iii) corrective feedback; that is, they were provided with the correct answer if their answer was incorrect. Participants then took a second test and a third a week later. The corrective feedback led to the best performance on both the second and third test, while the verification feedback was no more useful than receiving no feedback. Other studies conclude that it is generally better to provide learners with more elaborated feedback than just feedback in the form of 'correct' or 'not correct' (Bangert-Drowns et al., 1991; McKendree, 1990; Moreno, 2004; Pridemore, & Klein, 1995; Shute, 2008).

However, some studies do show a beneficial effect for verification feedback. Hanna (1976) conducted a study in which participants answered multiple-choice questions on science, mathematics, and social studies and then received (i) no feedback, (ii) verification feedback, or (iii) answer-until-correct feedback (something I refer to as 'trial-&-error', see section "*Trial-&-error*"). The results show verification feedback tended to be sufficient for high-performing learners, who were able to deduce the correct answer when informed that their answer was incorrect. Low-performing learners, on the other hand, were less likely to deduce the correct answer when informed that their answer was wrong. This group benefitted more from the answer-until-correct feedback.

Marsh, Lozito, Umanath, Bjork, and Bjork (2012) compared the effects of (i) no feedback, (ii) corrective feedback, and (iii) verification feedback. A total of 48 learners answered a series of general knowledge multiple-choice questions and took a test immediately the feedback was received with a second test after two days. The verification feedback was more useful than no feedback for improving on the final test, but corrective feedback was the most useful overall.

It appears that the benefit of verification feedback depends on the type of test as well as learners' ability level. The tests in Hanna's (1976) study were multiple-choice (though this was also the case for Pridemore and Klein (1995), and Marsh et al. (2012)). With such a test, indication of an incorrect choice tells more compared to a free-recall test; after all, it is possible to exclude at least one of the answers.

When it comes to learners' ability level, the study by Hanna (1976) showed that high-performing learners receiving verification feedback were likelier than low-performing learners to deduce the correct answer. Similarly, Schwartz et al. (2016) pointed out that for learners who already have basic knowledge within a specific area, simple verification feedback regarding whether an answer is correct or not, or whether a certain choice is adequate or not, can be useful. Very knowledgeable learners completing a familiar task may only need verification feedback. However, for novices, the sweet spot resides in more informative feedback.

Finally, the benefit of verification feedback depends on what other forms of feedback it is being compared to. Compared to no feedback, it at least provides a hint that the learner is heading in the right direction.

Trial-&-error feedback

One problem with verification feedback is that it is often accompanied by an opportunity to use trial-&-error. In principle, the learner could keep entering one answer after the other, without the need to put any thought into it. If the teacher only gets to see the final correct answer, s/he will have no information about how the student got there. In *Josefins skolvärld* (English: "*Josephine's School World*", Figure 3), the learner is supposed to click the number corresponding to the number of ladybugs in the picture. No matter how many times the learner clicks the wrong answer, s/he can continue until s/he gets the correct answer. Since each correct answer is rewarded by a point, there is no way to tell from looking at the scores whether a learner solved the task on the first, third, or twelfth trial.

Figure 3. An example of verification feedback where it is possible to use a trial-&-error strategy; upon a mistake, the learner is provided with a "wrong" buzz and can then try again for as many times as needed

If systematic trial-&-error helps a learner move forward with a low cost of time and effort, it is not hard to understand why the strategy can be the learner's first choice. From the perspective of the teacher though, trial-&-error behavior usually falls under the heading of 'gaming the system' defined by Baker et al. (2006, pp. 392-393) as "attempting to succeed in an educational environment by exploiting properties of the system rather than by learning the material and trying to use that knowledge correctly".

Returning to *Josefins skolvärld*, the design offers an opportunity to systematically click each answer until the right one is clicked, and the learner scores a point. The trial-&-error strategy, in this case, is a low cost in terms of effort and time spent – and there is no decrease in scores for errors. In fact, it can take less time to finish a task by repeatedly clicking on the different alternative answers than by really taking one's time and thinking the answers trough. This is particularly true for learners who are about to learn the content in question and generally need more processing time (that is, learners who are in their proximal zone of development and clearly could benefit from some more instructions and help).

Of course, the degree to which trial-&-error pays off is related to the design of the app. I will distinguish between 'low-cost trial-&-error', 'risky trial-&-error', and 'time-consuming trial-&-error'.

Low-cost trial-&-error is described in the example above. The learner can click randomly without risk of losing lives or scores.

Risky trial-&-error involves situations with a limit on time or the number of trials ("lives") allowed. If the learner is unlucky, s/he will not move on in the game and will have lost time, 'lives', and perhaps a chance to reach a high score.

Third, *time-consuming trial-&-error* involves situations in which trial-&-error behavior is likely to be very time-consuming. Consider a learner who cannot read and write who must spell a word, with all the letters in the alphabet at his/her disposal. It is possible in principle to take each letter and try it out – but solving the task this way will take a very long time. Other tasks are impossible to solve using this strategy – at least, not without a great deal of luck. Say, that a learner is supposed to find the sum 23 + 42, with no alternatives presented. A systematic attempt with all the numbers from one upwards will be (almost) impossible without thinking the answer through.

The use of trial-&-error strategies has repeatedly been shown to correlate negatively with learning (Aleven & Koedinger, 2001; Baker, Corbett, & Koedinger, 2004; Baker et al., 2006; Baker, Roll, Corbett, & Koedinger, 2005; Walonoski & Heffernan, 2006). Even though these studies have been carried out with middle-school students, there is reason to believe that the same applies to younger children.

Baker, Walonoski, Heffernan, Roll, Corbett, and Koedinger (2008) describe an animated agent designed to reduce the incentive to game the system: e.g., through trial-&-error. When such behavior is detected, the agent displays increasing levels of displeasure. If the learner by chance arrives at the correct answer anyway, the agent gives him/her a set of supplementary exercises covering the material that s/he has just skipped over. The results show that gaming-the-system behavior decreased overall, while learners who persisted in using such strategies increased their learning through the supplemental exercises they were given.

Corrective feedback

We have already learned that corrective feedback can be more beneficial than verification feedback (Marsh et al., 2012; Pashler et al., 2005). In addition, Clariana (1990) compared the effects of verification feedback (where a trial-&-error strategy was possible) and corrective feedback on 32 low-performing learners. The results show that corrective feedback had a significantly greater effect on performance than verification feedback. Furthermore, Phye and Sanders (1994) showed how more specific feedback in the form of providing the correct answer improved the performance on a retention task compared to more general feedback.

Corrective feedback provides the learner with more information than negative verification feedback in that it also provides the correct answer. Providing the learner with the correct answer can happen immediately after an incorrect choice, or it can be delayed until the end of the session. An example of corrective feedback can be seen in *Minilobes*; here the learner is asked to find the lower-case letter "a" (Figure 4, left panel). When the learner clicks the erroneous letter (lower-case "c"), the app directly says "cee" and then shows the correct answer (Figure 4, right panel).

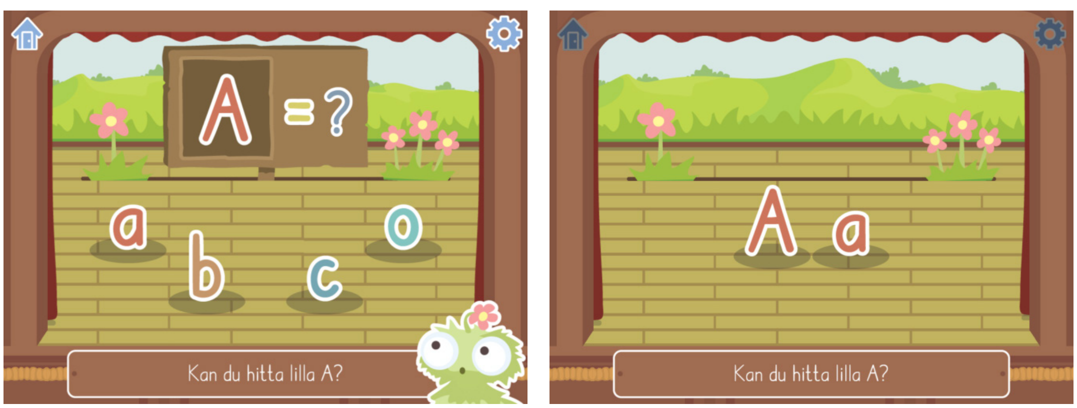

Figure 4. An example of corrective feedback: providing the learner with the correct answer after an incorrect one

The effects of corrective feedback again depend on what other forms of feedback are being compared, as well as the task and ability of the learner. Hattie and Timperley (2007) argue that this simpler feedback is most powerful when it addresses faulty interpretations rather than a total lack of understanding. Finn and Metcalfe (2010) found that corrective feedback seems to be beneficial for immediate testing, but not for delayed testing. Moreno and Mayer (2007), on the other hand, argue that novice learners learn better with explanatory feedback as compared to corrective feedback.

At the same time as corrective feedback provides the learner with more information than negative verification feedback, a potential drawback is that the learner, upon receiving the correct answer, just memorizes it without understanding. Rote learning is not a bad thing *per se*; but, in many cases, it is important first to have an understanding of what one is learning. Consider a child who learns that '2

× 3 = 6' but has little knowledge what the numbers mean. S/he does not understand that '2 × 3 = 3 + 3', which is the same as '2 + 2 + 2', and so on. Learning the multiplication table by heart only takes one so far.

Elaborated feedback

Elaborated feedback refers to any feedback that provides learners with more meaningful information. It comes in different forms and at different levels. Shute (2008) writes of elaborated feedback that it can choose to address the topic or the response; it can discuss specific errors, provide worked examples, give gentle guidance, or explain why a response was wrong and indicate the correct answer.

Elaborated feedback is generally associated with better learning (Bangert-Drowns et al., 1991; McKendree, 1990; Moreno, 2004; Pridemore & Klein, 1995; Shute, 2008). Further, Bangert-Drowns et al. (1991) argue that feedback is significantly more effective when it provides details of how to improve the answer instead of just indicating whether the learner´s work is correct or not.

Finn and Metcalfe (2010) conducted three experiments, comparing four situations: (i) corrective feedback, (ii) scaffolded feedback, (iii) answer-until-correct feedback, and (iv) minimal feedback, (participants were given one additional try when their first answer was wrong). The scaffolded feedback offered small hints guiding the learner toward the final answer, step by step: first providing the first letter in a target word, then the second letter, and so on.

The first experiment showed that the corrective and scaffolded feedback gave the best test scores upon immediate testing. In the second experiment participants not only were tested immediately but also after 30 minutes. The third experiment was exactly the same but with a delay of one day. Corrective feedback was best again with immediate testing, but scaffolded feedback gave best results for the delayed tests. Minimal feedback consistently produced the weakest results. The experiments show that when a learner just has a short time to correct an error, corrective feedback can be the best option, but that scaffolded feedback works best for long-term retention.

Facilitative feedback. Facilitative feedback means that the learner is offered a comment or suggestion to help him/her find the right solution. In *Läskod* (English: "Access code", Figure 5), the learner is to practice spelling. The task is to spell the word "*godis*" (English: "*candy*"). After a few mistakes, a question mark appears, and the student is allowed to see the correct spelling briefly before it disappears again. The problem with just showing the answer as a hint is that it allows the learner to copy-&-paste.

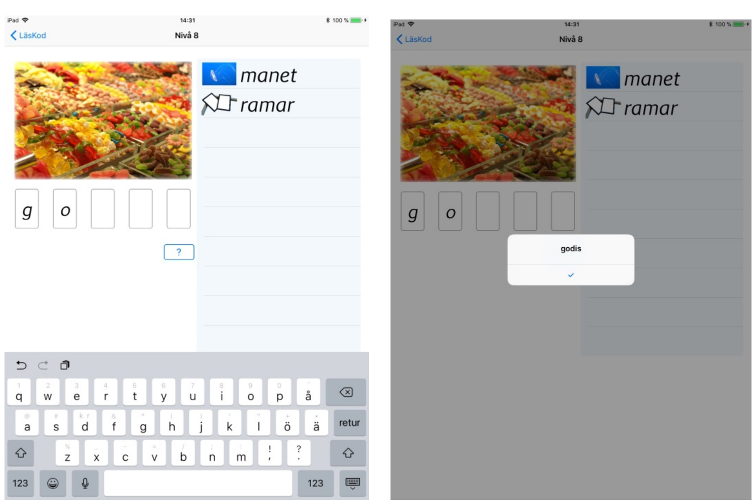

Figure 5. An example of facilitative feedback, showing the learner how to solve the task if s/he is struggling

Explanatory feedback

If facilitative feedback provides information on how the learner can solve the task, explanatory feedback provides more; for example, why the answer was correct or not. With explanatory feedback, the learner can build a deeper understanding of the task at hand and the foundations upon which to structure future tasks.

In the Swedish app *Särskrivning* (English: "*compound words written with a space between*", Figure 6) the learner is supposed to click on the expressions that are misspelled. In Swedish, 'compound words' are generally written without hyphen or space between them – and a misspelling can completely change the meaning, for example: "*sjuksköterska*" means "*nurse*" (in English), while "*sjuk sköterska*" changes meaning to "*sick nurse*" (in English). In the left panel, the learner mistakenly clicks on an expression that is spelled correctly "*gröna bönor*" (English: "*green beans*"): the program responds "*No, that was incorrect. The beans are described as green.*" In another task (right panel) the learner is asked to "*click on the right alternative*" for a common Swedish surname. The correct answer is "*Pettersson*". When the user clicks on "*Petters son*" (English: "*the son of Petter*"), the game responds "*Petter is a common Swedish first name – you just clicked on his son.*"

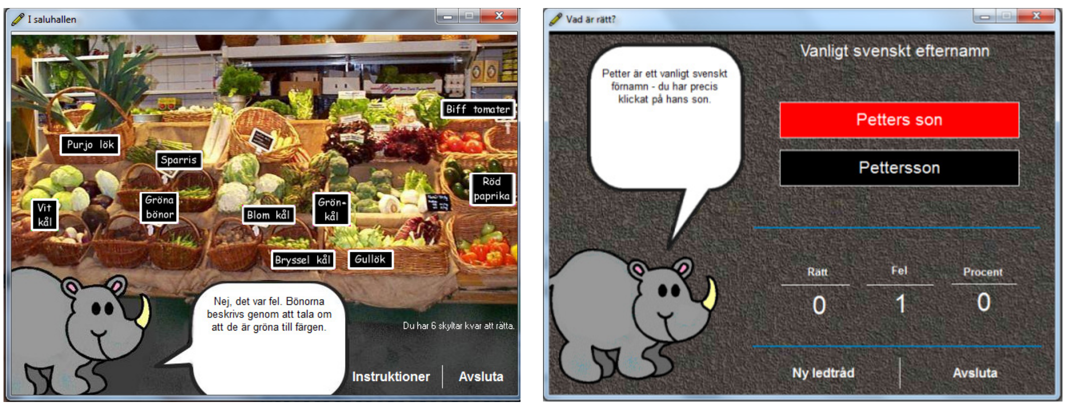

Figure 6. An example of explanatory feedback, where the learner receives information about why their answer was correct or incorrect

Moreno (2004) studied whether explanatory or corrective feedback worked best in a discovery-based learning environment, where novice learners were to design a plant capable of surviving under different weather conditions. Learners receiving explanatory feedback produced higher scores on a transfer test. Moreno argues that the explanatory feedback helped novices by decreasing their cognitive load, noting that benefits were found for cognitive but not affective outcomes; for example, motivation or interest.

Implication feedback

Actions and choices have consequences. If one miscalculates how many tablespoons of yeast one needs to add to dough, one will experience first-hand that the dough does not rise well. If a child is told to give apples to each of four horses but has only three and starts giving an apple to each, s/he discovers that one horse is left without an apple. The child learns more than just that his/her solution was incorrect: s/he may understand that three apples are too few for four horses but not *far* too few.

Such *implication feedback* (Blair, 2009), found in the mathematics game *Magical garden* (Figure 7), is meant to help preschoolers develop their understanding of number sense (Haake, 2018; Husain, Gulz, & Haake, 2015). Together with his/her teachable agent (a pedagogical agent whom the learner teaches at the same time as learning for him-/herself), the learner creates a magical garden by collecting water drops, which s/he receives by solving math problems. In one game, the learner is to help a hungry chameleon with weak eyesight catch ants. The learner can see if the chameleon's tongue reaches too low or too high (or catches the ant if the answer is correct).

Figure 7. An example of implication feedback: the learner sees that his/her answer is incorrect by watching the chameleon aim too high and so miss its food

Critter Corral (Figure 8), aims to help preschoolers develop concepts for the numbers one through ten. In one exercise, the learner's task is to fix a chair by choosing the correct leg size. If the learner chooses a leg that is too short or too long (or correct) this will be reflected in the game as well as stated by the speaker's voice. The thought behind this type of feedback is that the learner should develop a sense of magnitude and be provided with some guidance on how to revise their attempts (Blair, 2013). This in contrast with verification feedback, where the learner has to guess in which direction to go in order to fix a mistake. With implication feedback, in contrast, the learner is scaffolded by a hint.

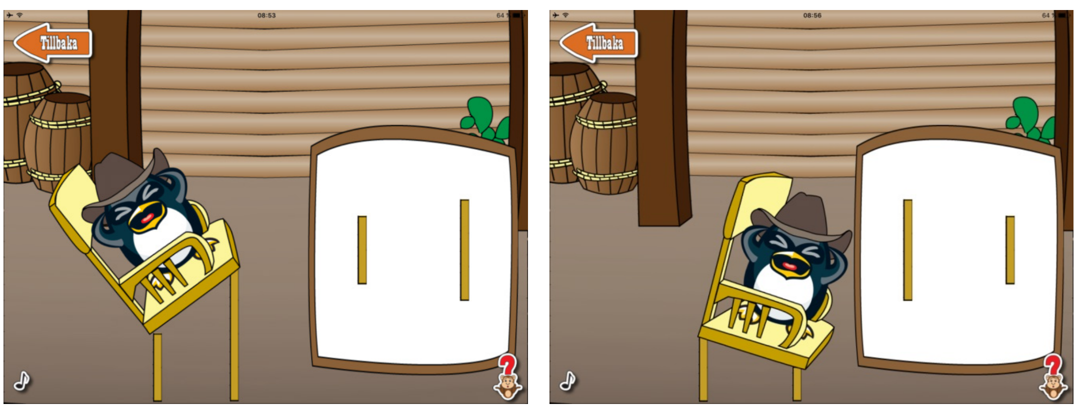

Figure 8. An example of implication feedback from *Critter Corral*: the learner is visually made aware of the implication of his/her answer by showing a too long or a too short chair leg

Feedback Focusing on the Learner

The kinds of feedback discussed so far all concern the task at hand and provided the learner with information about the task, the correctness of the task, and/or information about how to improve their solution to the task. In contrast, *encouraging feedback* and *result feedback* is information that concerns the learner rather than the task or how it can be solved.

Encouraging feedback

One role of feedback is to motivate the learner to continue with a task. *Encouraging feedback* is supposed to do this. Usually, such feedback is displayed as visual and/or auditory encouragements such as cheering, clapping, rising stars or balloons (or something happy) or via text or voice expressing how well the learner does (for examples of both kinds, see Figure 9).

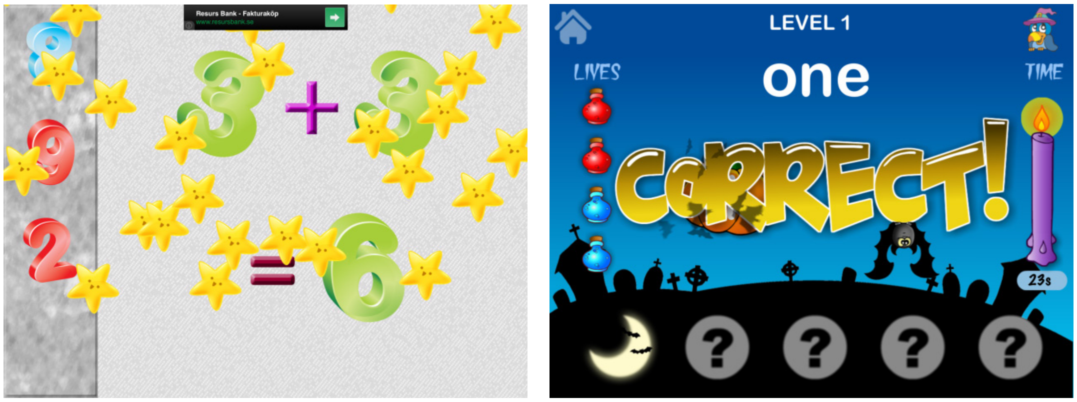

Figure 9. Two types of encouraging feedback

This type of feedback contains little (or no) task-related information, and the effects of it are rarely converted into more engagement, commitment to the learning goals, an enhanced self-efficacy or understanding of the task (Hattie & Timperley, 2007). Feedback about the self can even be seen as meaningless, and meta-analyses on teacher praise have found small, if any, associations with learner achievement (Kluger & DeNisi, 1996; Wilkinson, 1980).

The problem with this type of feedback is that it targets the learner as a person; for example, by saying "*good girl*" or "*you are brilliant*". It does not say anything about *what* the learner did well (and perhaps less well). It does not contain any information about the learner's effort involved in trying to solve the task or in managing to solve it more effectively. Hattie and Timperley (2007) point out that the highest effect sizes with respect to learning were found in studies that involved learners who received feedback about the task and how to solve the task more effectively. Praise, rewards and punishments were associated with much smaller effect sizes. Feedback about the self, such as "*you are a great learner*", cannot really, as the authors point out, help the learner to proceed in his/her learning.

This is not to be confused with how praise regarding achievement and learning can sometimes assist in enhancing self-efficacy, which may in turn influence achievement (Hattie & Timperley, 2007). Nicol and Macfarlane-Dick (2006) further argue that praising effort and strategic behavior leads to higher achievement compared to praising ability and intelligence. This is also supported by Black and Wiliam (1998), who recommend avoiding feedback that draws attention away from the task and towards self-esteem since this can have a negative effect on attitude and performance. Learners are much better served by praise for the efforts they invest in a task than by praise earned by their innate abilities (Dweck, 2000). Praising only children's intelligence can lead them to avoid tasks in which they could potentially learn something due to the fear of looking stupid or losing face (Dweck, 2000; Gunderson et al., 2013).

We should think twice before praising the learner (at least without thinking about what we are praising), which does not mean, however, that learners do not like to be praised – they most often do (Burnett, 2002; Elwell & Tiberio, 1994; Sharp, 1985).

Result feedback

A common form of feedback in school is results, such as a student's score or grade. This is also quite common in apps, and I will refer to this as *result feedback*. Just as with encouraging feedback, this type of feedback does not provide the learner with information about the learning process or a how the learner could improve. It is a mere evaluation of how well the learner has performed during, for instance, a game session. Notably, this kind of information can be misguiding for someone who looks only at the result and who may not, for instance, be able to tell whether the presented result is an outcome of a low-cost trial-&-error strategy.

If the goal is for a learner to evaluate his/her own progress, such feedback can be a good parameter to use, but when used as a tool for comparing the performance of different learners, it may lead to stress and negative feelings for some students (but for some who like to compete it might also be encouraging). When result feedback is used in order to compare learners, the focus turns more to the learner than to the difficulties in a task and efforts to improve. Simply put, receiving a grade or a result can risk making the learner focus on the wrong thing, and if no other feedback is given, a simple result does not tell the learners how they could improve. This, in turn, has been shown to have a negative effect on motivation (Butler, 1987; Craven, Marsh, & Debus, 1991; Harlen & Deakin Crick, 2003).

Feedback in Educational Software

There is not a large amount of research on the types of feedback used in educational software. However, there are some recent studies that review educational software more broadly, and some consider feedback, even when the concept as such is not used.

Cherner, Dix, and Lee (2014) put forward a framework for how to choose educational apps based on their purpose, content, and value. Some researchers have examined different categories of apps (Handal, El-Khoury, Campbell, & Cavanagh, 2013; Highfield & Goodwin, 2013). Larkin (2015) evaluated apps for mathematics, analyzing how many of them provided the learner with conceptual knowledge (i.e. information that involves understanding related to the meaning of mathematics), procedural knowledge (the ability to follow a set of sequential steps to solve mathematical tasks) and declarative knowledge (information that the learner retrieves from memory without hesitation). Highfield and Goodwin (2013) reviewed the pedagogical content within the most popular apps in Australia, UK, and the USA, and found that 74% of all apps had elements of 'drill and practice', tasks that require minimal cognitive investment on behalf of the learner. These types of tasks usually require minimal cognitive investment and frequently use extrinsic rewards. From this review, it can be concluded that more apps need to be developed that also focus on children's ability to develop as self-regulatory learners, who do not only memorize things by heart without understanding. Hirsh-Pasek et al. (2015) offer a way to define the potential educational impact of current and future apps. Along the same line, Sjödén (2017) evaluated what factors are important when evaluating an educational app, and here feedback is mentioned as one of the cornerstones.

The present study focuses on the types of feedback that are represented in apps commonly used in Swedish schools today. The focus lies on the information regarding the task that different types of feedback provided to learners and whether or not these different types of feedback are beneficial for learning. According to previous research, more elaborated feedback provides learners with more information regarding their misconception and thereby provides them with more solid ground on which they can build their knowledge. Based on findings from Highfield and Goldwin (2013), Blair (2013), Sjödén (2017), and my own experience with apps, I predicted that few of the apps would contain elaborated feedback – which according to the literature would be most appropriate in order to enhance learning.

Method

I distributed an email to primary schools around Sweden asking them to send a reply regarding which apps they used. The email was distributed to approximately 40 schools, and 14 of them replied. The answering schools were distributed from Luleå in the north of Sweden to Ystad in the south. The target software was apps used for children at their school (i.e. grades 1-6).

NUMBER OF APPS REVIEWED

In total, I received the names of 164 different apps, of which several were used at more than one school. Since I had not asked for apps targeting a certain subject, the received apps targeted various subjects such as mathematics, Swedish, programming, learning the clock, biology, and geography.

All apps were read about in App Store or on the internet before being analyzed. This resulted in some apps being removed from the analysis since they were considered to be non-educational, explorative, supporting apps, too complex apps, or apps that could not be found.

In total, I removed 61 apps out of the 164; 25 because I considered them to be non-educational in that they did not cover any subject in the curricula (for example the camera, *Gmail*, the calculator, *iMovie*, etc.), 7 others that are used only for exploration (an example of this category was an app in which the learner could practice how to do different geometrical shapes with digitalized rubber bands; at no time were the students given any feedback on what shape they had built or how their shape could be transformed into another shape, etc.). Six apps were categorized as supporting tools for the learner and/or teacher (such as an app translating a Swedish word into English). In addition, there were 23 apps that I found too complex to evaluate, since they were parts of a larger learning environment or because they could not be found in *App Store*.

After removing 61 apps from the sample, I was left with 103 apps to analyze. Several of these apps contained subgames, which in this study are treated as individual apps since they touch upon different subjects or have a different gaming idea. So, from the original 103 apps, a further 169 apps (i.e. subgames) were added to the analysis, making it 272 apps. When reviewing these, 29 were removed, since they did not contain any activity where the learner could do something categorized as wrong.

For example, in *Bugs and bubbles* (Figure 10, left panel) the learner's task is to collect all green dots by tilting the tablet in different positions. The learner may miss a dot at one trial, but then s/he can just tilt the tablet so that the ball takes another round (preferably past the green dot). In *Siffermix 1* (Figure 10, right panel) the learner is supposed to click any number; the number is then represented by a set of objects.

Figure 10. Examples of apps that were not included in the review, since the learner cannot do anything categorized as wrong

One additional app was removed from the sample since it did not provide the learner with any feedback. Having thus removed 30 subgames, I was left with 242 apps (including subgames) to review. Some apps were available as both a free and a paid version; in 13 cases there was both a free and a paid/commercial version, and in 8 of these the free version was used (this was in cases where the judgment was made that a payment would not bring anything extra).

I played each app for as many times as it took to grasp the gist of it and establish what types of feedback were present. While playing, I was consciously trying to make as many mistakes as possible to see in what way the app would provide me with feedback and possibly guide me towards the correct answer. In addition, I also tried different trial-&-error strategies to see if any of my categorized strategies could be used. Approximate gaming time was between 20 minutes and 1 hour per app.

The purpose of this review was to study types of feedback in apps and to discuss whether the types of feedback found support learning or not. The different types of feedback analyzed can be seen in Figure 1.

This review only concerns feedback and does not look at other factors that could influence learning. This means that an app, in this review, might be portrayed as less satisfactory concerning the feedback it provides or how feedback is provided, while it may still have other, more positive features. For example, an app might help the learner to visually represent a number that s/he is supposed to calculate (Figure 11). Here, the learner is supposed to add "3 + 2", but instead of only showing the numbers, which a learner at a given stage might find abstract and have a hard time grasping and finding meaningful, the number is also represented with cookies in different colors. Visualizing a number can make it easier for some learners to do the calculation, and instead of abstract numbers, there are concrete objects to handle.

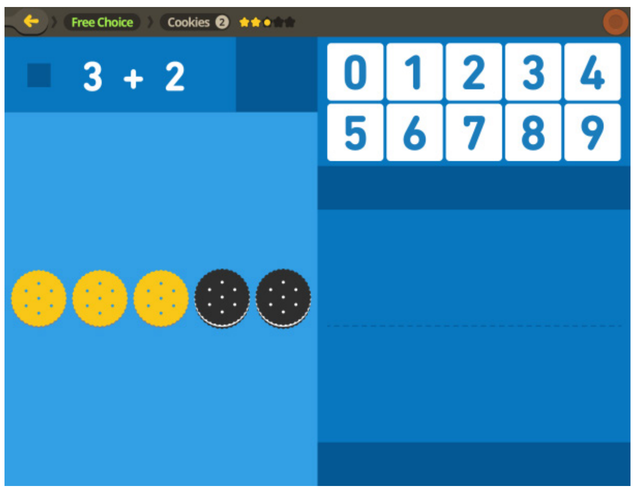

Figure 11. An example of visual representation in the app *Todo math – cookies*

RESULTS AND DISCUSSION

Out of the 242 apps reviewed, 187 (77%) contained verification feedback only; that is, the app provided the learner with information as to whether an answer was correct or incorrect. Twenty-five apps (10%) contained corrective feedback in that the right answer was displayed after the player had provided an incorrect answer. Further, 30 apps (12%) contained information in the form of elaborated feedback (facilitative, explanatory, or implication feedback), see Figure 12.

These results confirm my prediction that few apps would contain elaborated feedback, and the percentage was as low as 12%. From what we know from the literature, more elaborated feedback is preferable if learning is to be supported (Bangert-Drowns et al., 1991; McKendree, 1990; Moreno, 2004; Pridemore & Klein, 1995; Shute, 2008). Thus, most of the apps reviewed do not fulfil that requirement.

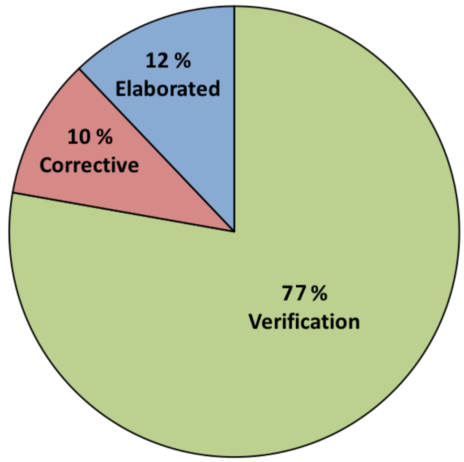

Figure 12. The proportion of apps providing each of three types of feedback

Verification Feedback

A potential pitfall with verification feedback is that it may encourage different trial-&-error strategies, which the learner can use in order to complete a task. I therefore analyzed the possibilities of using low-cost trial-&-error, risky trial-&-error, and time-consuming trial-&-error (Figure 13).

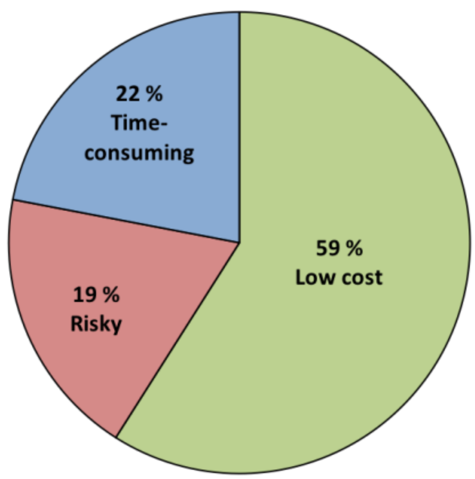

Figure 13. Verification feedback and its division in different trial-&-error strategies

Low-cost trial-&-error feedback

Fifty-nine percent of the apps that only included verification feedback was designed in such a way that low-cost trial-&-error strategies could be used. That is, more than half of the reviewed apps made it possible for the learner to get all answers correct without actually having to pay attention to the task at hand. Typical examples are provided in Figure 14. In the app *ABC-klubben*, the task is to drag the card that starts with the letter M to the empty square (Figure 14, left panel). When the wrong card is drawn (here, the picture of a fire), the card simply returns to its starting position, and a 'negative' sound can be heard. The learner can then try again for as many times as s/he wants until s/he chooses the correct card, and the app confirms the correct answer by saying "*monster*".

Another typical example can be seen in *Lola's Mattetåg* (Figure 14, right panel), in which the learner is supposed to click on the number three. When clicking the incorrect answer, Lola (the panda) shakes her head, and the erroneously clicked answer is highlighted with a red ring. Again, the learner can try for as many times as s/he wants until s/he gets it right.

Figure 14. Examples of apps in which the learner can use a low-cost trial-&-error strategy; the learner can click randomly without risk of losing lives or scores

As already mentioned, low-cost trial-&-error strategies are not beneficial for learning (Aleven & Koedinger, 2001; Baker et al., 2004, 2005, 2006; Walenoski & Heffernan, 2006). The learner could very well be thinking about other things and not the task at hand, but still gain a high score. Clicking without paying attention to the task is not likely to lead to any good learning. However, from an outsider's perspective, a high score and a fast response time indicate that the learner is good at the task, something that might be false.

Risky trial-&-error feedback

In 19% of the apps that contain only verification feedback, the learner can apply a risky trial-&-error strategy. This means that with a little luck the learner may provide the correct answer by chance. But there are also other elements, such as time spent and 'game lives' lost, that need to be taken into account. The chances to get a high score or a fast time by just guessing are lower than in low-cost trial-&-error, and if the learner replies by chance every time, the chances of him/her getting a high score or moving on to the next level are slim, since every incorrect answer is 'punished' in some way such as, for example, losing a life or points or not levelling up.

The left panel in Figure 15 (*Happi läser*) shows an example in which the learner needs six watermelons (as can be seen at the bottom of the left panel) in order to move on to the next level. Typing in an incorrect answer provides a 'negative' sound and the pictures disappear and are replaced by new ones. That is, clicking an incorrect picture will not give the learner any disadvantages, but at the same time, s/he will not be able to move on to the next level unless s/he provides six correct answers.

Another example can be seen in the app *Math king* (Figure 15, right panel), in which the learner is supposed to sum up the numbers represented by the fingers. The learner only has three 'lives', and each time an incorrect answer is provided, s/he also loses score, which can be used to climb a 'career ladder'. Here the learner has more to lose compared to low-cost trial-&-error apps (at least if the learner wants to progress within the app).

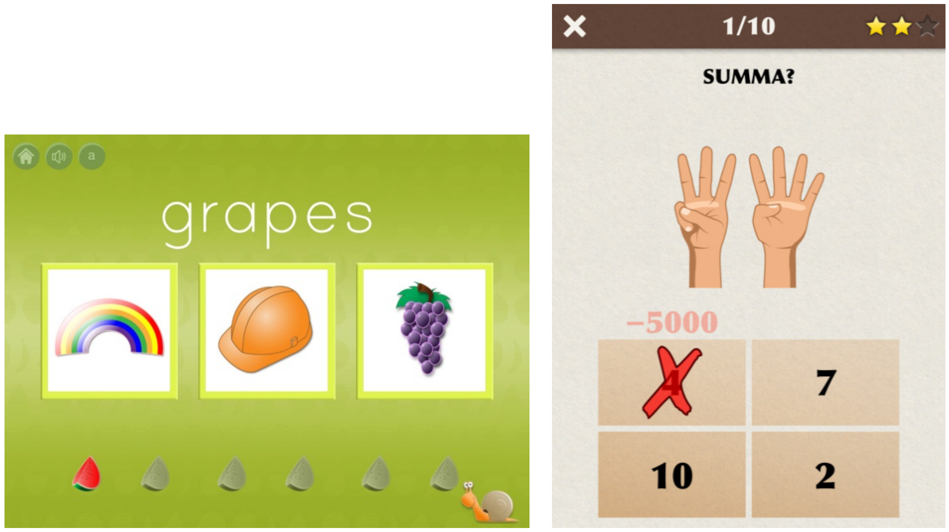

Figure 15. Two examples of apps where a risky trial-&-error strategy can be used; if clicking on an erroneous answer, s/he runs the risk of not moving on in the game, losing lives or the chance of reaching a high score

Time-consuming trial-&-error feedback

The last 22% of the apps that contain only verification feedback were apps in which the task could be solved by using a time-consuming trial-&-error technique and also apps in which it is (practically) impossible to solve the task in this way if the learner has no idea of how to solve it. Examples of this can be seen in Figure 16.

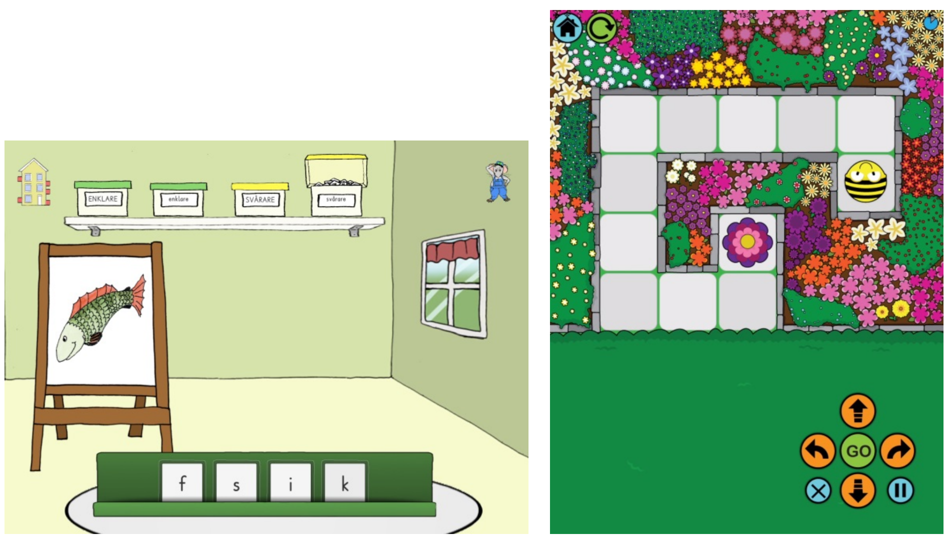

Figure 16. Two examples of apps where a time-consuming trial-&-error strategy can be used

In *Bornholmslek – Bygga ord* (Figure 16, left panel), it is possible to find the correct answer by using the strategy of trying out every possible combination of letters provided. Here the learner is supposed to spell "*fisk*" (English: "*fish*"). Even though it is time-consuming, it is not impossible to try different combinations of letters until it is correct. The learner is also provided with feedback in the form of sounds telling him/her how a certain letter is pronounced. By using this information, it is possible to find the correct solution without knowing it from the beginning.

In the app *Bee-Bot* (Figure 16, right panel) it is, on the other hand, practically impossible to find the correct solution, unless the learner has an idea from the beginning about how to solve the task. The learner is here supposed to guide the bee to the flower by using programming commands. In addition, the difficulty of the problems increases considerably, in that the learner has to keep every command in their working memory – there is no visualization of commands already ordered.

In tasks like this, the learner would probably be helped by, first of all, receiving some type of command tracing, so that they would not have to keep their commands in their head. In addition, if the learner could also trace their commands in combination with the bee's path, it would make it more visible for the learner where a possible error in their coding had occurred. What is also troublesome is that the bee always starts from the beginning of the commands; if not, it would be possible for the learner to take it one step at a time as they can in another app called *Lightbot* (Figure 17).

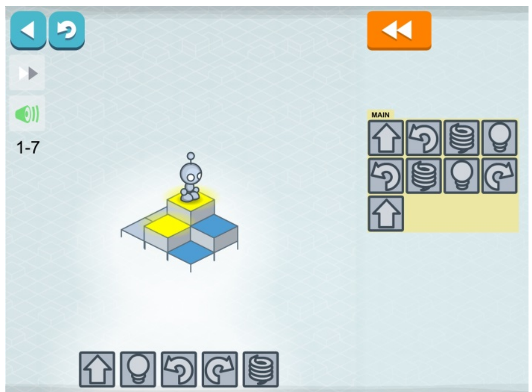

Figure 17. An example of visual feedback showing the learner which commands s/he has pressed

In Lightbot, the learner receives better scaffolding to trace their programming since the commands are visualized and don't need to be kept in one's working memory. It can still be difficult for the learner to see exactly where their programming went wrong. If it was possible to slow the robot down even further, as well as making the robot walk at the same time as the corresponding command was lit up, the feedback would be even clearer, at least at the beginning, when the task might still be new and challenging.

Common for all apps that contain only verification feedback is that a learner who does not know the correct answer from the beginning can solve the tasks, yet still be left with knowledge gaps. They can also be stuck on a task without knowing how to fix it, since no further feedback is provided. For example, in Lightbot again, if the learner cannot figure out by testing how to make the robot light all the blue boxes, this can cause frustration, since there is no help available for each step the learner should take in order to reach the goal.

If the aim of the app is to teach something and for a learner to develop knowledge, skills, or understanding s/he did not have before, there should be some feedback helping the learner if s/he needs it. If the goal of an app, instead, is to test knowledge, understanding, or skills that one believes are in place, the demands on the app are different.

Corrective Feedback

Twenty-five apps contained more information than only verification feedback in that they also provided the *correct* answer when the learner typed in an *incorrect* one. Two examples are shown in Figure 18.

In the app *Math bingo* (Figure 18, left panel), the correct answer, in this case "1 + 1 = 2", is shown after the learner has provided an incorrect answer. The learner then receives a new task to solve, here

"7 + 4 = ?" In the app *Geoexpert*, the learner is shown a flag (right panel in Figure 18, top left corner) and the name of the corresponding country. The learner's task is to find the country among those marked on the map and click it. After two incorrect answers, the correct country is circled. An addition in this app is that if the learner clicks an incorrect country (such as clicking at Brazil), the app displays Brazil's flag as well as types the name 'Brazil'. Hereby the app provides the learner with information that s/he may use later.

Figure 18. Two examples of corrective feedback, in which the learner is provided with the correct answer after an incorrect one

Another example can be seen in the app *Räkneapan* (Figure 19) in which all numbers that the learner replied incorrectly to are summarized at the end of the game. By getting all the incorrect answers summarized at the end, the learner is given the opportunity of studying them further.

Figure 19. Corrective feedback shown at the end of the game

Although not many explanations are provided in the analyzed apps, the learner is not left with a complete blank as to what was wrong with their answer, since they are provided with the correct one. By being presented with the correct answer, they gain some information that may be used for learning.

ELABORATED FEEDBACK

Twenty-nine out of 242 apps (12%) contained some type of elaborated feedback: facilitative feedback, explanatory feedback, or implication feedback.

Facilitative feedback

Most of the elaborated feedback has the form of being facilitating, providing the learner with some type of hints on how to solve the task at hand. Out of the 29 apps that contained elaborated feedback, I categorized 23 as providing facilitative feedback. An example of an adequate or useful hint provided to a learner can be seen in Figure 20, showing the app *Mattebageriet 2*. If the learner is not able to solve the task "*12 + 20*" (Figure 20, left panel) there is a light bulb in the upper left corner, which can be clicked, and the app then provides a hint that asks the learner to count "*How many single cookies are there on the plate*?" (Figure 20, right panel). This hint provides the learner with information that is useful for solving the task and also tells him/her where to start. If the learner has little experience in solving these types of tasks, just knowing where to start can be a problem and the task can seem overwhelming. Feedback that guides a learner towards the correct answer can be helpful for many. This app provides the learner with further hints if s/he doesn't know how to move forward.

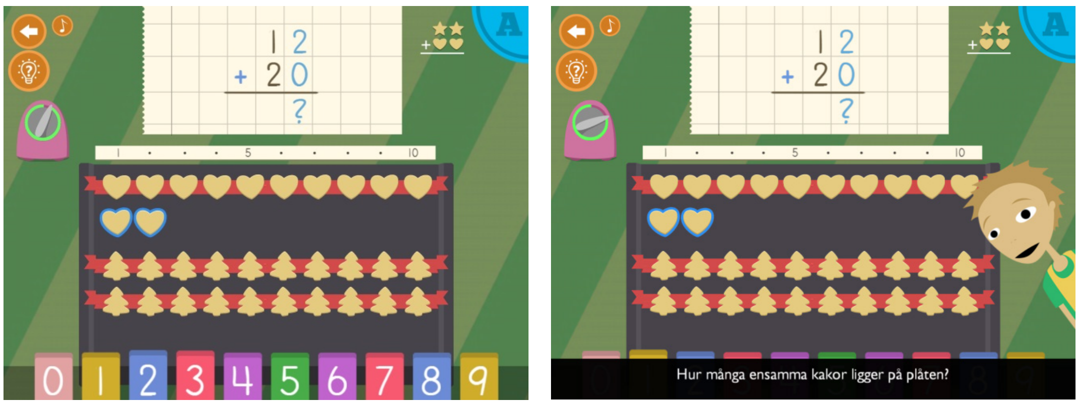

Figure 20. An example of facilitative feedback where the learner is provided with hints guiding him/her towards the correct answer

Another example, from the app *Bokstavspussel*, can be seen to the left in Figure 21 in which the learner is supposed to spell the word "giraff" (English: "giraffe"). After a first incorrect try one letter is revealed, after two incorrect tries a second letter is revealed, and so forth.

This type of feedback provides a small part of the solution in order to help the learner spell the word correctly. It can potentially be problematic if a learner mindlessly drags whichever letter to a random place just to learn where to put one letter. After seven tries (in this case with "*giraff*") the answer will be shown and could just be copied. In cases when learners actually do make an effort and try to spell the word correctly, the feedback provided can, however, provide adequate support for learning. It can be compared to the beneficial effects of scaffolding feedback that Finn and Metcalfe (2010) found in their studies.

It is more likely that the strategy of copy-&-paste, which is not desirable from a learning perspective, is applied by learners who use *Happi stavar* (Figure 21, right panel). This game aims to let learners practice spelling with cross puzzles. The learner can try on his/her own, but if s/he gets stuck, there is a lightbulb in the upper left corner, which makes the correct spelling appear in the background. Again, for a learner who doesn't know how to spell a certain word or for a learner who doesn't want to make an effort, the task can easily be solved by just copying the correct answer (after clicking the light bulb).

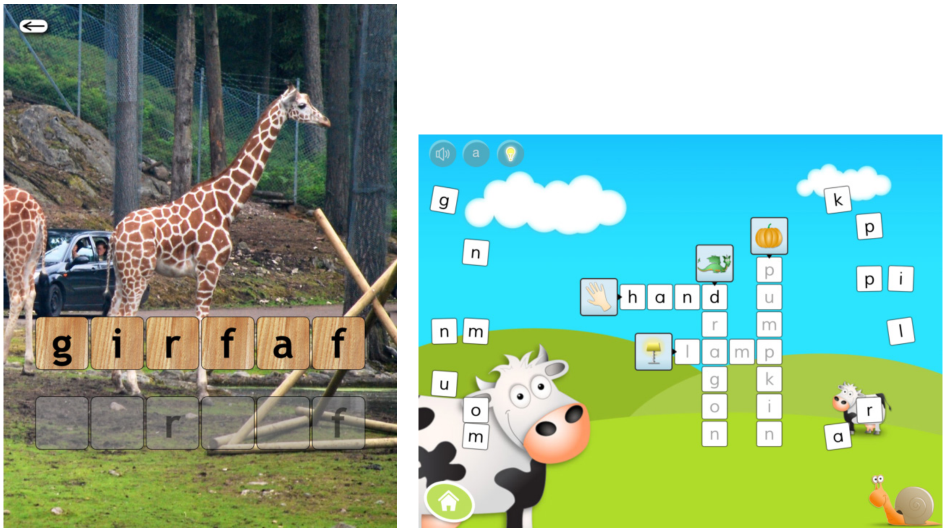

Figure 21. Two examples of facilitative feedback in which the learner could choose a copy-&-paste strategy

Another type of hint can be seen in Figure 22. Here the learner's task is to find the numbers that equal 10. After a few incorrect tries a hint appears in the upper left corner, providing the learner with an example of such numbers ("8 + 2 = 10"). Providing the learner with hints like this may remind him/her what s/he is supposed to do, and also what one possible solution could look like. Showing examples of performance can make explicit to the learner what is required; in addition, it can define a standard (Orsmond, Merry, & Reiling, 2002). However, in this game incorrect answers can also originate from the fact that the learner does not remember where certain numbers are situated; in other words, it is not merely a mathematical task, but also a memory task.

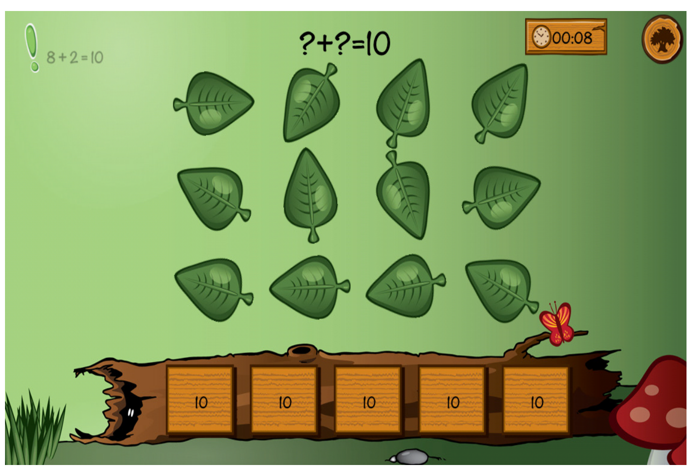

Figure 22. An example of facilitative feedback, showing a type example

Further, there are apps that do provide good facilitative feedback but still suffer from certain problems. In the app *Todo math – light it up* the learner is supposed to practice counting. As demonstrated in Figure 23, the learner is supposed to solve "10 + ? = 15". The starting position shows ten blocks on the number line, but also a yellow triangle showing the final sum. In general, this is a very good example of facilitative feedback, since the learner receives support in the form of visual representations from the number line, and they get to see implications of their answers (implication feedback) when adding too few (Figure 23, right panel) or too many (Figure 23, left panel) boxes to the line. A problem, though, is that everything is shown from the beginning. This means that the learner does not have time to think the numbers through. Already from the beginning, the starting position and the end position are shown, and the learner only has to fill in the blanks. According to Bangert-Drowns et al. (1991), this type of information (where the learner will not have time to verbalize an answer of their own) can even have negative effects on learning.

On the other hand, the provided hints are good and would probably work great for learners who are struggling with these types of tasks. An alternative could be to let the learner have a go without the number line and blocks, and these could be added one at a time when the learner needs them.

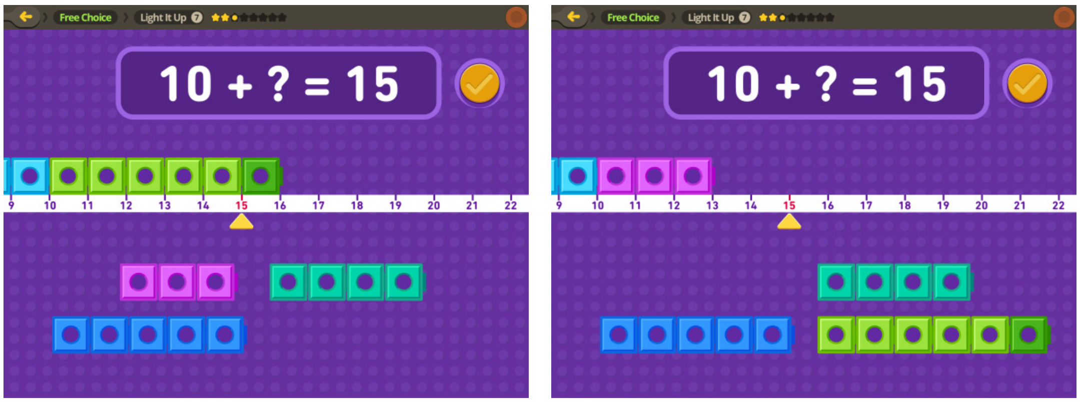

Figure 23. Example from *Todo math*, which guides the learner towards the task; a possible problem is that all information is given from the start, without giving the learner time to think the task through

A similar problem is found in *Motion math – fractions*. Here the learner is supposed to tilt the tablet to make a bubble bounce at different fractions (¼ in the example in Figure 24). After a first incorrect bounce, an arrow appears, showing in which direction the learner should move the bubble. After a second missed bounce, lines appear (Figure 24, second panel), displaying a visual representation of the whole number. After one more mistake, the app displays the incorrect fraction the learner bounced the bubble on (Figure 24, third panel). As one last hint, the app displays an arrow showing the learner the correct answer. Then the learner can try again for as many times as s/he wants.

Again, the facilitative feedback is well designed and aims at helping the learner reach the correct answer. However, the problem is that the response has to be so quick that learners may have a hard time reaching the correct answer. The ball is bouncing at a predetermined rate, and if you are a slow thinker or are having trouble with how to tilt the device, you will not have the time to make a correct bounce. Removing the time factor or being able to choose at what rate the learner wants it to bounce, could possibly make the task easier. Or why bounce at all?

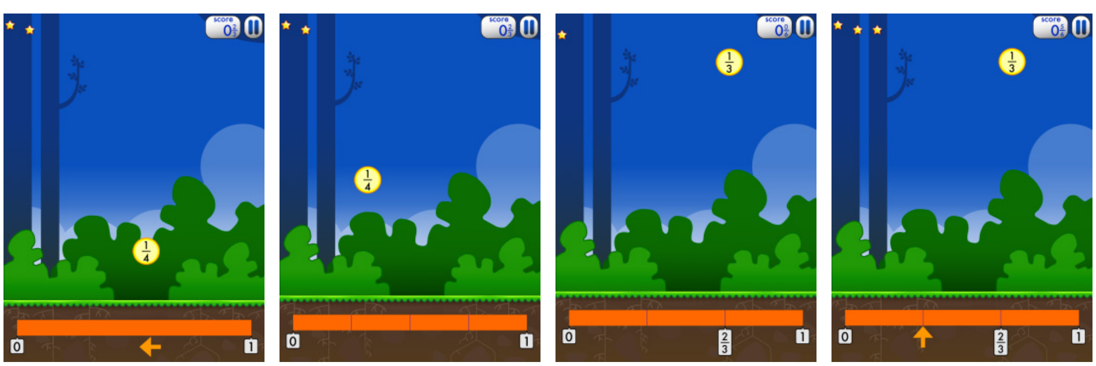

Figure 24. Example from *Motion math – fractions*, which guides the learner towards the correct answer; a possible problem is that that the hints are shown too fast, so that the student won't have the time to think the answer through

A better example of facilitative feedback can be found in a subgame in the app *Vektor*, where the learner is asked to represent the number in the grey box (here number "5", see Figure 25). The learner starts out with a timeline, where the numbers 0, 5 and 10 are visually shown. If s/he doesn't succeed in three trials, additional facilitative feedback is provided in the form of more numbers shown on the line (Figure 25, upper right panel). If the learner still doesn't succeed, the app displays additional hints in the form of an arrow showing the correct answer (Figure 25).

Likewise, if the learner does succeed with the task, the hints are removed one at a time so that no numbers are shown in the end. One may argue that there is a possibility of applying a strategy of not succeeding on purpose, in order to be able to copy-&-paste the answer in the end, yet the learner has to try three times before a hint is displayed, and when s/he succeeds with three trials in a row, the hints are again removed.

Figure 25. A good example of facilitative feedback in that it provides the learner with hints on how to solve the task; the hints are removed when the learner succeeds with the task at hand

Explanatory feedback

Only two apps out of the entire sample of 242 apps fulfil the criteria of providing explanatory feedback. In *Zcooly affären 2* (Figure 26), the learner takes the role of a cashier with the task of providing the customer with the articles asked for. The learner also has to charge the customer the correct amount by putting money in the cash machine. When the learner does something wrong, the app tells him/her what was wrong, for instance, that the customer has received the wrong items or has been charged too much or too little.

In the second example (Figure 27), the app provides both explanatory and implication feedback. The learner owns a bakery, in which s/he bakes cupcakes to sell in order to make money. Then with more money s/he can buy more ingredients and bake more cupcakes, and so forth. In the upper left panel in Figure 27, the customer requests x number of cupcakes. After having delivered the cupcakes, the learner is provided with information regarding incomes versus expenditures, and s/he can see the implication of his/her income and expenses in the form of earned money (Figure 27, upper right panel).

Figure 26. An example from the app *Zcooly-affären 2*, in which the learner should provide the customer with the correct groceries and charge the customer. The learner is provided with explanatory feedback when s/he does something incorrectly.

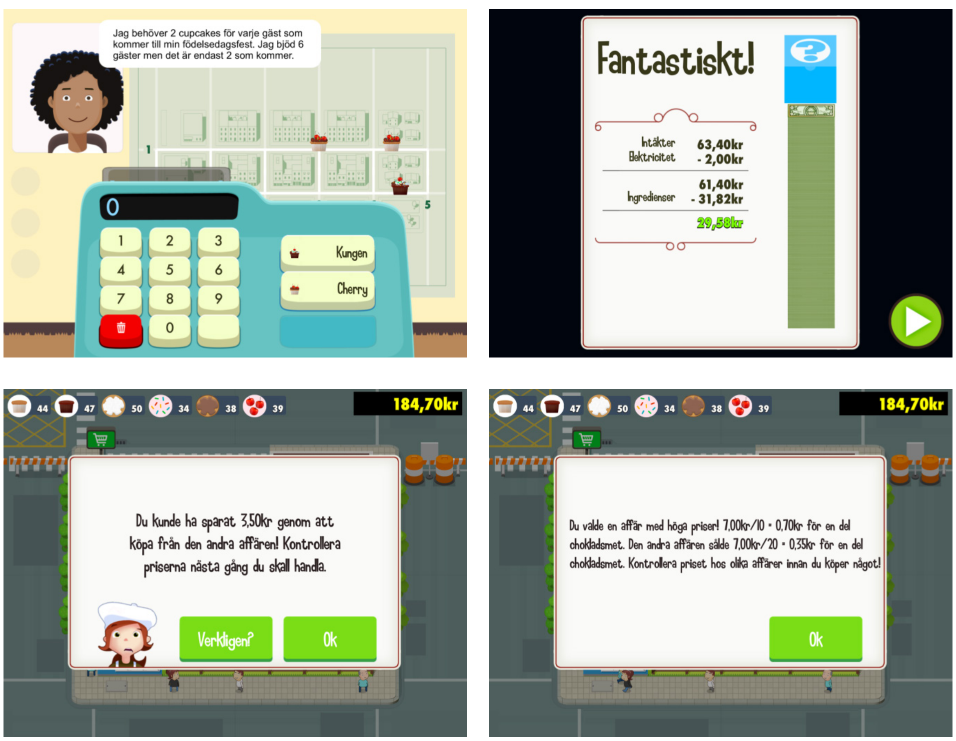

Figure 27. An example of both implication and explanatory feedback. Upper left: a customer telling the learner how many cupcakes s/he would like to buy. Upper right: the learner is shown the implications of his/her income and expenses. Bottom left and right: explanations to the learner regarding their purchase and why this was not the best purchase.

Further on, there are two stores from which the learner can buy his/her ingredients for the cupcakes. After such shopping the learner receives information about his/her purchase, telling him/her whether s/he could have saved money by going to the other store, or if s/he made the best available purchase (Figure 27, bottom left panel where it says "*You could have saved 3,50 SEK by shopping in the other store! Check the prices next time you are out shopping*"). The app also provides the learner with an opportunity to find out more about his/her purchase (in this case overly expensive) by clicking "*Really*?" instead of "*OK*". Further explanations regarding his/her purchase are then provided (Figure 27, bottom right panel "You chose a store with high prices! 7.00 SEK/10 = 0.70 SEK for one batch chocolate dough. The other store sold one batch of chocolate dough for 7.00 SEK/20 = 0.35 SEK. Compare prices in different stores before you buy anything!").

Here are two examples in which the student could have benefited from some explanatory feedback. In *Farm factor* (Figure 28) the learner´s task is to fill in the number s/he thinks corresponds to the number of radishes in the basket. When the learner fills in "5 × 3" (Figure 28, left panel), the app tells her that this is incorrect. S/he also receives a hint telling him/her that "*The multiplication symbol* × *means 'groups of'*." In the left part of the left panel of Figure 28, it says: "*There are 3 groups of 5*.", but this is easily missed if the learner is just concerned with calculating the number. Also, the hint tells the learner nothing of how his/her answer "*5* × *3*" is the same as the requested "*3* × *5*" from a mathematical point of view. More informative feedback explaining to the learner in what sense it is adequate to equate "*3* × *5*" and "*5* × *3*" and in what sense it is not, which relates to what kind of answer is searched for in this task, ought to be provided. Without this, the learner might in the worst case believe that "*5* × *3*" does not equal "*3* × *5*".

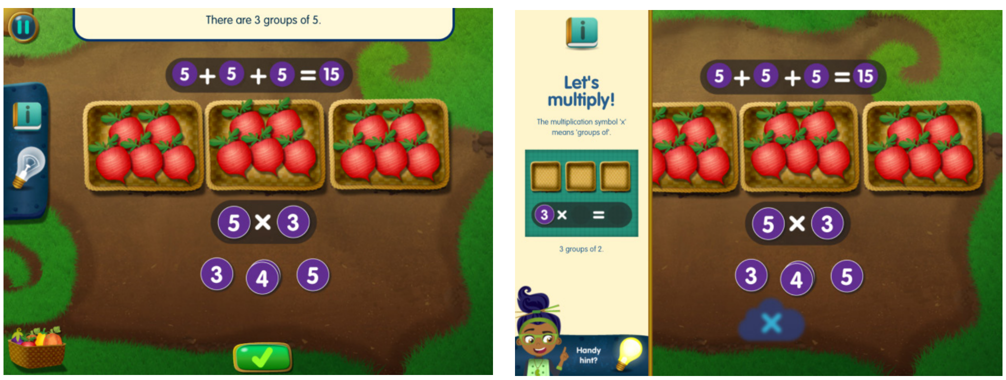

Figure 28. Example of a math problem in which the learner would have benefited from some explanatory feedback telling him/her *in what way* his/her answer was incorrect

In a similar example (Figure 29), the learner is supposed to spell "*ambulans*" (English: "*ambulance*"). When pressing "a" as the first letter, s/he receives feedback telling him/her that this is wrong. But why is it wrong? Well, there are two a's in *ambulans*, and the student picked the 'wrong' one.

Figure 29. Example of a situation in which the learner makes an error but receives no explanation as to why this (assumable correct) answer was incorrect

Implication feedback

Three additional apps contained implication feedback. In all these apps, the learner is supposed to solve tasks involving a scale and weights. In Figure 30, for example, the learner is supposed to arrange weights and choose the correct rule of arithmetic in order for the scale to balance at zero. When choosing an incorrect weight or arithmetic rulemaking, the scale leans in some direction (as can be seen in Figure 30), so the learner can visually see that his/her answer is incorrect. At the same time, s/he also gets an implication as to what side is heavier and thereby adds up to a larger sum.

Figure 30. An example from *Todo math – number balance* in which the student receives visual implications that here the answer is incorrect

ENCOURAGING FEEDBACK

Encouraging feedback often occurs in combination with some of the other types of feedback. In my review, I have found two types of encouraging feedback. One is encouragement in the form of applause, cheering, balloons and stars, or other displays that appear after a task has been completed. In the other type, encouragement comes in the form of spoken or written utterances evaluating the learner's performance, such as "*good work*", "*perfect*", or "*amazing, you did it*". Around half of all apps in this review (55%) contain some type of encouraging feedback. In turn, 53% of these make use of spoken or written utterances that comment on how well the learners perform, and the remaining 47% make use of balloons, cheering, etc. (see Figure 31 for two examples).

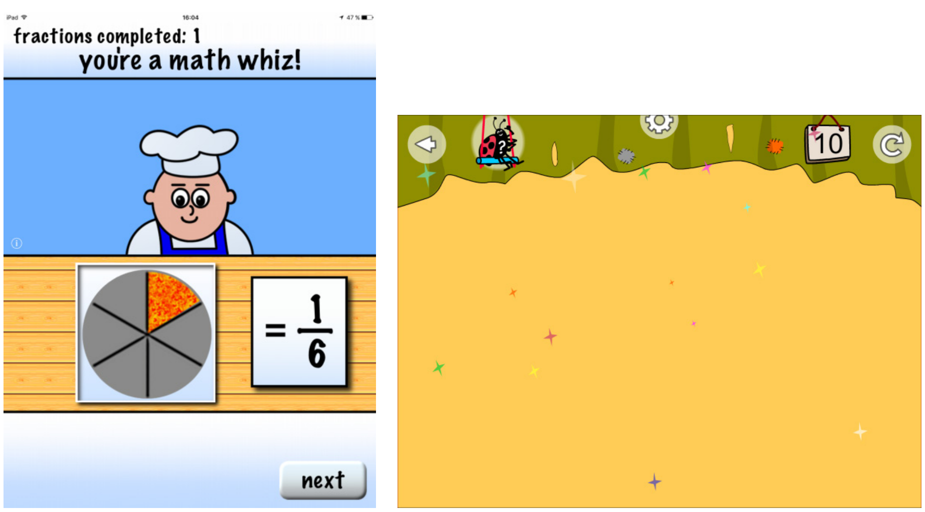

Figure 31. Two examples of encouraging feedback: to the left, encouragement towards the learner, and to the right, encouragement in the form of stars falling after a completed task

As mentioned above, encouraging feedback rarely leads to higher performance or to higher self-efficacy (Hattie & Timperley, 2007). Nonetheless, learners do like to be praised (Burnett, 2002; Elwell & Tiberio, 1994; Sharp, 1985). More than half of the apps in the present review contained various kinds of encouragement and praise. Most likely, such feedback does not boost the learner's performance. On the other hand, if learners like it and it does not *decrease* performance, it should be fine.

Yet, the picture is more complex and worth digging into. This type of feedback can have a negative effect when it becomes obvious to the learner that there is no relation between the feedback and what actually goes on. For example, as illustrated in Figure 32, the learner (me) correctly answered 21 questions out of 80 (making it 59 incorrect answers), but the sign still says that I "*did a great job*" and that "*this was awesome*".

Even though it may not be wise to say, "*this was not so very good*", the response "*this was awesome*" might make the learner question the apps credibility, since most likely the learner has a sense of how well (or not so well) s/he has done on the task. Another known effect is that a learner risks thinking along the lines: *"You don't think much of me if you say this was awesome.*" or "*So, no one expects more than this from me.*"

Figure 32. An example of encouraging feedback that despite the "not so good" result of having 21 out of 80 questions correct encourages the learner by writing "*Well done!*", "*Really good!*" and "*Awesome!!!*" on the board to the right: encouragement that could be perceived as unrealistic

Almost all encouraging feedback in the entire range of apps was delivered after a *successful* trial. It is worrying how extremely unusual it was that an app contained any form of encouragement when the learner *did not* succeed with the task. Only 9 apps encouraged the learner to try again. In Figure 33, the learner is encouraged to continue with the task by hearing things like "*not completely right, try again*" or "*there is a picture that fits the sound better, click the mouse with striped pants to hear the sound again*". This at least acknowledges that the learner clicked an incorrect answer and encourages him/her to go for another round. No app gave the learner any encouragement or praise for his/her effort, saying that the learner is doing a great job putting so much effort into the task or that s/he has fought well when doing something wrong.

According to Hattie and Timperley (2007) and Dweck (2000), comments targeting the learner's intelligence and/or ability are problematic, since they turn the focus to the person and not the task; something that students can perceive as threatening. From a learning perspective, it is preferable to comment on the efforts and/or steer the focus towards the task.

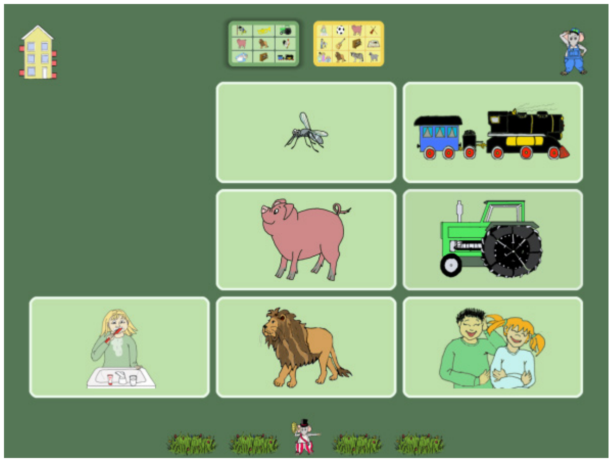

Figure 33. In the app *Bornholmslek – Ljud* the learner is encouraged to continue with the task when clicking the incorrect picture.

RESULT FEEDBACK

When it comes to result feedback, 94 apps out of 242 provided the learner with information presenting him/her results. This is information that can be used to compare to other learners or between own results; for instance, to see whether one is making progress. Examples of result feedback can be seen in Figure 34, where the learner receives the result and his/her personal high score. Often these results are received in combination with some type of encouragement (see Figure 34 right panel) where the learner receives the comment "*CLOSE ENOUGH! You scored 5 out of 10 […]*".

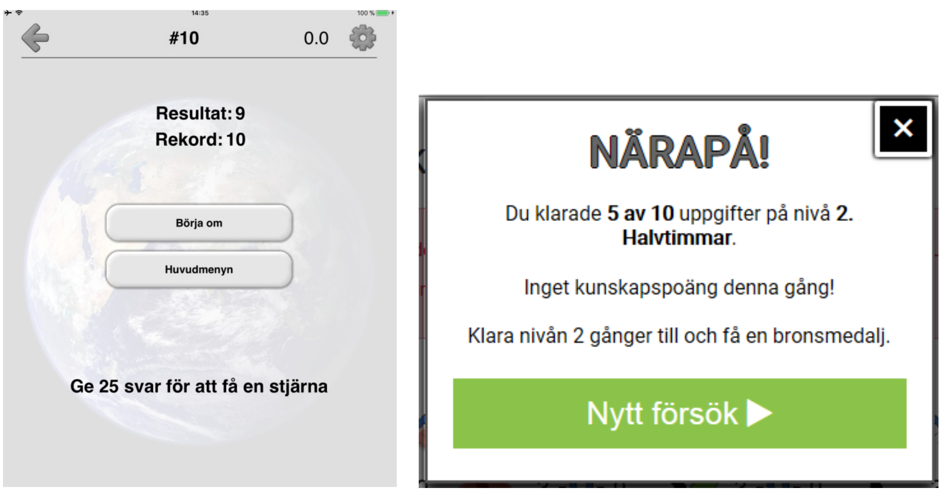

Figure 34. Two examples of result feedback. Left panel shows: "*Results: 9; High score: 10*". Right panel shows: "*Close enough! You scored 5 out of 10 tasks at level 2.*"

Eighteen apps showing result feedback presented the number of correct answers needed in order to move on to the next level. In Figure 35, the learner needs five correct answers in a row in order to finish. The number of incorrect answers is not displayed, so a regular comparison to other results cannot be made. This could be positive, in that it is impossible for a learner to compare him-/herself to others, which can potentially cause stress and negative feelings. To be noted is that this type of feedback can also be positive for some students, who see the results as encouragement and as an incentive to try harder.

From a teacher's perspective, this feedback can be problematic, since the only result the teacher will ever see is the number of correct answers, and s/he will not know what types of questions the learner struggled with. Another example is when only the correct answers are summarized and displayed. There are several apps designed in this way.

Figure 35. Example of result feedback, only shown after the goal is reached, saying "*Congratulations! 5 correct answers in a row!*"

SUMMARY AND CONCLUSION

As predicted, a majority of the apps do not provide learners with elaborated feedback. In fact, only 29 out of 242 (twelve percent) provide anything more than only verification or corrective feedback. From a learning point of view, this is disappointing. If we look at the literature, most research emphasizes the importance of elaborated feedback for learning (Bangert-Drowns et al., 1991; McKendree, 1990; Moreno, 2004; Pridemore & Klein, 1995; Shute, 2008). In other words, most feedback provided in apps today is more suited for testing situations and does not provide the learner with more elaborated information from which they can correct mistakes and misconception or build new knowledge.

VERIFICATION FEEDBACK

Verification feedback was the predominant type of feedback, and in 78% off all apps this was the only kind of feedback provided. With this type of feedback, the learner will know whether their answer is correct or not, as it provides some sort of guidance. For a learner with prior task knowledge, such guidance can be sufficient as support towards the correct answer after they have made a mistake. But for a learner who does not have such prior task knowledge, just knowing whether their answer was correct or not will not help much.

Similarly, if the learner believes that his/her answer is correct, whereas the feedback says this is not the case, this may cause frustration and helplessness. Being told that you are wrong without any further guidance telling *why* or *how* can be problematic. Providing the learner with the correct answer (corrective feedback) will at least provide him/her with some information – but involves other disadvantages. Being presented with the solution instead of being allowed to actively come up with it yourself is often less powerful in terms of understanding and remembering.

A problem associated with verification feedback is that it encourages using trial-&-error strategies. In this review, I categorized three different trial-&-error strategies. 'Low-cost trial-&-error ', in which the learner can move forward at a low cost in terms of time and effort, is the type of verification feedback that is the most problematic from a learning perspective. It is very common that apps allow learners to use this strategy – in this review, 59% of all the apps provided verification feedback only. A low-cost trial-&-error strategy allows the learner just to click different answers until the correct one is hit, and there are no consequences when clicking an incorrect one.

That it is *possible* to use a low-cost trial-&-error strategy is not a problem with learners who actually try to solve the task and who are making an effort. But with learners who only want to 'get by' and would rather *not* make an effort, this possibility is troublesome. Research has shown that so-called 'gaming the system' is negatively related to learning (Aleven & Koedinger, 2001; Baker et al., 2004, 2005, 2006; Walonoski & Heffernan, 2006). Although this strategy is not used by all learners, the learners not using it are, in general, not the ones we have to worry about.

Similarly, if the learner does not know the correct answer, pure guessing – which is possible in apps that allow low-cost trial-&-error strategies – can let the learner finish the task with a good score and in good time. From a perspective from the outside, it might seem as if the learner knows what s/he is doing, whereas in fact, little learning has occurred.

Pure guessing, which can be used in apps that allow for low-cost trial-&-error strategies, is constrained in an app that only allows for what I call risky trial-&-error. In this case, if the learner uses the strategy of pure guessing, there is a cost in terms of 'lives', scores, or levelling. That is, every mistake the learner makes costs him/her; for example, a life or several points. Low-cost trial-&-error strategies are possible in 19% of all apps in the review that contain verification feedback only.

It is important to point out here that there is no ideal type of feedback, which will always work best for all learners in all situations. It is also not the case that verification feedback is always inferior to other kinds of feedback or a bad design choice. As already mentioned, verification feedback can be just what a high-performing learner with sufficient prior knowledge needs to work on a given task and to learn from it. Also, if the purpose of an app is to test or evaluate knowledge or skills, verification feedback is adequate. The mismatch may arise if the app is advertised as an app that supports learning.

Corrective Feedback

Extending verification feedback by adding corrective feedback provides the learner with somewhat more information: at least they won't have to wonder what the correct answer should be. Potentially, they may also use the provided correct answers for further learning. In this review, 10% of the 242 apps contained corrective feedback. That is, they provide the learner with the correct answer when s/he proposes an incorrect one. Previous research has shown that corrective feedback is more beneficial for learning compared to verification feedback (Marsh et al., 2012; Pashler et al., 2005).

There is a caveat: if the learner only memorizes the provided correct answers without reflecting on them and, if appropriate, trying to understand why a particular answer is correct, the resulting learning may be shallow. A piece of information learned by heart, with no knowledge of how and in what situations to use it, will not lead anywhere. Knowing *that* this was the right answer does not equal knowing *why* this was the correct answer. Again, learners with adequate prior knowledge are more likely to figure out why an answer is correct, whereas for students who are less knowledgeable this will be harder or impossible.

Elaborated Feedback

When it comes to *elaborated feedback*, 23 out of the 29 apps that provide more than only verification or corrective feedback contained *facilitative feedback*. This refers to feedback that provides some kind of hint on how to solve or proceed with the task. This can be a good way of guiding the learner towards the correct answer if s/he is stuck with a task. Yet, in some cases, such as, for example, in *Happi stavar* (Figure 21, right panel), it opens up for the use of a copy-&-paste strategy if the learner wants to avoid making an effort (or completely mistrusts his/her own abilities to learn and to solve tasks).

Explicit explanations as to why a certain answer is correct or not were only provided by two apps. In *Motion math cupcakes* the learner is provided feedback on whether and why s/he made the best purchase s/he could when buying ingredients for his/her cupcakes. This information pinpoints an important feature, namely that comparing prices between the two stores could save him/her some money in future purchases. Even though this app contains somewhat more complex tasks, so that it may be more obvious that explanatory feedback is an adequate feature, other apps could very well benefit from it as well. For example, in a spelling app, if a learner is spelling the word "*träd*" (English: "*tree*") with two ä's, the app could tell the learner that this was almost correct, but that in the Swedish language we seldom use two vowels in a row, with a few exceptions like "*zoo*" and "*leende*" (English: "*smile*").

Another way to provide the learner with more information about how to reach the correct answer is to provide implication feedback. This is also the way we often encounter feedback in our everyday life. Four of the apps reviewed contained this type of feedback, and in three of them, the task concerned math and a balancing scale. Even though it is encouraging to see that this type of feedback is prevalent in this specific domain, it should also be possible to provide this kind of feedback for many other types of tasks (cf. *Critter Corral*: Blair, 2013). This could, for example, be applied in *Lolas mattetåg* (Figure 14) where the learner sometimes has to solve a math task by adding two numbers. Such an addition could be as follows. If the learner answers correctly, Lola's train will reach the train station, but if the learner proposes a sum that is too large, the train moves past the station, and if the learner proposes a sum that is too small, the train stops before the station.

Encouraging Feedback and Result Feedback

Encouraging feedback almost always comes in combination with some other type of feedback, and 133 out of 242 apps (55%) contained either encouraging feedback in the form of cheering, balloons, etc., or messages in text or voice saying that the learner is awesome, is doing perfectly or very well. In all apps that make use of written expressions, the feedback targets the learner and not the task. In other words, it is the learner who is praised for being smart, doing great, etc. The focus is on the child – not on the task. Addressing intelligence and/or ability in this manner has not been shown to be beneficial for learning. On the contrary, it can make the learner focus on the wrong things and lead them to avoid future tasks in which they risk failing (Dweck, 2000; Gunderson et al., 2013).

Yet, encouragements and praise are often appreciated by learners, and they can indeed be useful. The recommendation is also not to eliminate encouragement and praise but to shift the focus from the learner to the task or to the effort that the learner puts into the task. Adding encouraging feedback telling the learner that s/he is making progress, making a good effort, does not seem to give up easily, etc. should not be an impossible design task for app designers. Overall, learners can use some encouragement when they have made a mistake but continue working on the task, and not only when they have already finished the task. Only 9 out of 133 apps encouraged the learner in some way to continue. This could be done more often. From my experiences of talking to teachers in schools and preschools, they are well aware of the drawbacks of praise that focuses on the person ("*you are really bright*", "you are very good in math", "*oh, you are smart*"). Instead, they praise and encourage with a focus on the task or what has been produced ("*this is very well done*", "*I like how you solved this*", "*this essay is very well written*"). In addition, they all agree on the importance of encouraging effort and providing feedback during the working and learning process. There is a striking mismatch between teachers' views on encouraging feedback and the implementation of encouraging feedback in educational apps.

When it comes to result feedback, 39% of 242 apps presented results that a learner can use to compare him-/herself to others or use as a measure of his/her own progress. For competitive learners, this can be a good way to motivate themselves to continue and try harder, but for learners who do not appreciate competition or have low beliefs in their own ability, it can instead be stressful. In most cases, this kind of feedback tells the learner nothing about in what respects they need to practice more. They will not know which questions they answered wrong or which topics they did less well at. This type of feedback therefore rarely leads to increased learning, but it can be used by some learners as a measure of when they have to work harder.

CONCLUSION

The use of digital apps is increasing in schools today, and in order for them to be useful as learning devices, and not only testing devices, they need to provide feedback that is more informative than only telling whether a choice or answer was correct or incorrect. One advantage with technology is that it offers an opportunity to provide all learners with the same or individualized feedback at the same time, and it is up to the designers to make the most out of this, just as it is up to them to reduce the opportunities for trial-&-error; in particular, low-cost trial-&-error.

However, reading an introductory text for an app will often not reveal whether the app is indeed a learning – and not a testing – device. It is not forbidden to use the term 'supports learning' in a text that describes an app, even though no learning scientist would approve. While working on this review, I read through all available information texts. Only 4 out of 99 (texts were not found for all 103 apps) stated that the app in question was designed 'to test or evaluate skills and knowledge', whereas many more are suited for testing purposes – but not for learning purposes. The only way to know if a certain educational app matches the purpose you have – whether as a teacher or a parent – is to play it yourself and try to make as many mistakes as you can.

A tentative conclusion on the basis of this review is that many educational app designers view a learner as someone just waiting to be informed whether an answer or a choice was correct or not. This kind of feedback corresponds to a behavioristic approach comparable to instrumental conditioning by means of reinforcement. In essence, most apps miss the opportunity of treating the learner as an active and constructive being who would benefit from more nuanced feedback.

REFERENCES

Aleven, V., & Koedinger, K. R. (2001). Investigations into help seeking and learning with a cognitive tutor. In R. Luckin (Ed.), *Papers of the AIED 2001 Workshop on Help Provision and Help Seeking in Interactive Learning Environments* (pp. 47-58).

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1-18. https://doi.org/10.1037/a0021017

Baker, R. S., Corbett, A. T., & Koedinger, K. R. (2004). Detecting student misuse of intelligent tutoring systems. In J. C. Lester, R. M. Vicari, & F. Paraguaçu (Eds.), *7th International Conference on Intelligent Tutoring Systems* (pp. 531-540). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-540-30139-4_50

Baker, R. S., Corbett, A. T., Koedinger, K. R., Evenson, S., Roll, I., Wagner, A. Z., ..., & Beck, J. E. (2006). Adapting to when learners game an intelligent tutoring system. In M. Ikeda, K. D. Ashley, & T.-W. Chan (Eds.), *LNCS, Proceedings of the 8th International Conference on Intelligent Tutoring Systems* (Vol. 4053, pp. 392-401). Berlin, Germany: Springer.

Baker, R. S., Roll, I., Corbett, A. T., & Koedinger, K. R. (2005, May). Do performance goals lead learners to game the system? In C.-K. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Eds.), *Proceedings of Frontiers in Artificial Intelligence and Applications* (Vol. 125, pp. 57-64). Amsterdam, The Netherlands: IOS Press.

Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., & Koedinger, K. (2008). Why learners engage in "gaming the system" behavior in interactive learning environments. *Journal of Interactive Learning Research*, *19*(2), 185.

Bangert-Drowns, R. L., Kulik, C. L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61(2), 213-238.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education, *5*(1), 7-74. https://doi.org/10.1080/0969595980050102

Blair, K. P. (2009). *The neglected importance of feedback perception in learning: An analysis of children and adults' uptake of quantitative feedback in a mathematics simulation environment* (Doctoral Dissertation, Stanford University). Retrieved from https://eric.ed.gov/?id=ED532821

Blair, K. P. (2013). Learning in critter corral: Evaluating three kinds of feedback in a preschool math app. *Proceedings of the 12th International Conference on Interaction Design and Children* (pp. 372-375). New York, NY: ACM. https://doi.org/10.1145/2485760.2485814

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn*. Washington, DC: National Academy Press.

Brockbank, A., & McGill, I. (1998). *Facilitating reflective learning in higher education*. Bristol, PA: Taylor & Francis.

Burnett, P. C. (2002). Teacher praise and feedback and learners' perceptions of the classroom environment. *Educational Psychology*, *22*(1), 5-16. https://doi.org/10.1080/01443410120101215

Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, *79*(4), 474-482. https://doi.org/10.1037/0022-0663.79.4.474

Cherner, T., Dix, J., & Lee, C. (2014). Cleaning up that mess: A framework for classifying educational apps. *Contemporary Issues in Technology and Teacher Education*, *14*(2), 158-193.

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73-105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Clariana, R. B. (1990). A comparison of answer until correct feedback and knowledge of correct response feedback under two conditions of contextualization. *Journal of Computer-Based Instruction*, *17*(4), 125-129.

Craven, R. G., Marsh, H. W., & Debus, R. L. (1991). Effects of internally focused feedback and attributional feedback on enhancement of academic self-concept. *Journal of Educational Psychology*, *83*(1), 17-27. https://doi.org/10.1037/0022-0663.83.1.17

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development.* Oxford, UK: Psychology Press.

Elwell, W. C., & Tiberio, J. (1994). Teacher praise: What learners want. *Journal of Instructional Psychology*, *21*(4), 322-328.

Finn, B., & Metcalfe, J. (2010). Scaffolding feedback to maximize long-term error correction. *Memory & Cognition*, *38*(7), 951-961. https://doi.org/10.3758/MC.38.7.951

Fox, B. A. (1991). Cognitive and interactional aspects of correction in tutoring. In P. Goodyear (Ed.), *Teaching knowledge and intelligent tutoring* (pp. 149-172). Norwood, NJ: Ablex.

Gibbons, P. (2002). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom* (pp. 110-119). Portsmouth, NH: Heinemann.

Gunderson, E. A., Gripshover, S. J., Romero, C., Dweck, C. S., Goldin-Meadow, S., & Levine, S. C. (2013). Parent praise to 1-to 3-year-olds predicts children's motivational frameworks 5 years later. *Child Development*, *84*(5), 1526-1541. https://doi.org/10.1111/cdev.12064

Haake, M. (2018). No child left behind, nor singled out – reasons for combining adaptive instruction and inclusive pedagogy in early math software. Manuscript submitted for publication.

Handal, B., El-Khoury, J., Campbell, C., & Cavanagh, M. (2013, September). A framework for categorising mobile applications in mathematics education. *Proceedings of the Australian Conference on Science and Mathematics Education, Canberra, Australia,* 142-147. Retrieved from https://researchonline.nd.edu.au/edu_conference/70/

Hanna, G. S. (1976). Effects of total and partial feedback in multiple-choice testing upon learning. *The Journal of Educational Research*, 69(5), 202-205. https://doi.org/10.1080/00220671.1976.10884873

Harlen, W., & Deakin Crick, R. (2003). Testing and motivation for learning. *Assessment in Education: Principles, Policy & Practice*, *10*(2), 169-207. https://doi.org/10.1080/0969594032000121270

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112. https://doi.org/10.3102/003465430298487

Higgins, R., Hartley, P., & Skelton, A. (2001). Getting the message across: the problem of communicating assessment feedback. Teaching in Higher Education, 6(2), 269-274. https://doi.org/10.1080/13562510120045230

Highfield, K., & Goodwin, K. (2013). Apps for mathematics learning: A review of 'educational'apps from the iTunes App Store. In V. Steinle, L. Ball, & C. Bardini (Eds.), *Proceedings of the 36th Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 378-385). Melbourne, Australia: MERGA.

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in "educational" apps: Lessons from the science of learning. *Psychological Science in the Public Interest*, *16*(1), 3-34. https://doi.org/10.1177/1529100615569721

Husain, L., Gulz, A., & Haake, M. (2015). Supporting early math: Rationales and requirements for high quality software. *The Journal of Computers in Mathematics and Science Teaching*, *34*(4), 409-429.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75-86. https://doi.org/10.1207/s15326985ep4102_1

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254-284. https://doi.org/10.1037/0033-2909.119.2.254

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, *1*(4), 279-308. https://doi.org/10.1007/BF01320096

Lepper, M. R., Aspinwall, L., Mumme, D., & Chabay, R. W. (1990). Self-perception and social perception processes in tutoring: Subtle social control strategies of expert tutors. In J. M. Olson, M. P. Zanna, & C. P. Herman (Eds.), *Self-inference processes: The Ontario Symposium* (Vol. 6, pp. 242-257). Hillsdale, NJ: Lawrence Erlbaum Associates.

Larkin, K. (2015). "An app! An app! My kingdom for an app": An 18-month quest to determine whether apps support mathematical knowledge building. In T. Lowrie & R. Jorgensen (Eds.), *Digital games and mathematics learning* (pp. 251-276). Dordrecht, The Netherlands: Springer.

Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, *23*(2), 157-172. https://doi.org/10.1080/03075079812331380364

Marsh, E. J., Lozito, J. P., Umanath, S., Bjork, E. L., & Bjork, R. A. (2012). Using verification feedback to correct errors made on a multiple-choice test. *Memory*, *20*(6), 645-653. https://doi.org/10.1080/09658211.2012.684882

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, *59*(1), 14-19. https://doi.org/10.1037/0003-066X.59.1.14

McKendree, J. (1990). Effective feedback content for tutoring complex skills. *Human-Computer Interaction*, *5*(4), 381-413. https://doi.org/10.1207/s15327051hci0504_2

Moreno, R. (2004). Decreasing cognitive load for novice learners: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science*, *32*(1-2), 99-113. https://doi.org/10.1023/B:TRUC.0000021811.66966.1d

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, *19*(3), 309-326. https://doi.org/10.1007/s10648-007-9047-2

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199-218. https://doi.org/10.1080/03075070600572090

Orsmond, P., Merry, S., & Reiling, K. (2002). The use of exemplars and formative feedback when using learner derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education*, *27*(4), 309-323. https://doi.org/10.1080/0260293022000001337

Orsmond, P., Merry, S., & Reiling, K. (2005). Biology students' utilization of tutors' formative feedback: A qualitative interview study. *Assessment & Evaluation in Higher Education*, *30*(4), 369-386. https://doi.org/10.1080/02602930500099177

Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 3-8. https://doi.org/10.1037/0278-7393.31.1.3

Phye, G. D., & Sanders, C. E. (1994). Advice and feedback: Elements of practice for problem solving. *Contemporary Educational Psychology*, *19*(3), 286-301. https://doi.org/10.1006/ceps.1994.1022

Pridemore, D. R., & Klein, J. D. (1995). Control of practice and level of feedback in computer-based instruction. *Contemporary Educational Psychology*, *20*(4), 444-450. https://doi.org/10.1006/ceps.1995.1030

Schwartz, D. L., Tsang, J. M., & Blair, K. P. (2016). *ABC's of how we learn – 26 scientifically proven approaches, how they work, and when to use them.* New York, NY: W. W. Norton & Company.

Segedy, J. R., Kinnebrew, J. S., & Biswas, G. (2013). The effect of contextualized conversational feedback in a complex open-ended learning environment. *Educational Technology Research and Development*, *61*(1), 71-89. https://doi.org/10.1007/s11423-012-9275-0

Sharp, P. (1985). Behaviour modification in the secondary school: A survey of students' attitudes to rewards and praise. *Behavioral Approaches with Children*, *9*, 109-112.

Shute, V. J (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153-189. https://doi.org/10.3102/0034654307313795

Sjödén, B. (2017, November). What teachers should ask of educational software: Identifying the integral digital values. *Proceedings of 10th annual International Conference of Education, Research and Innovation, Seville, Spain,* 6491-6500. https://doi.org/10.21125/iceri.2017.1673

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295-312. https://doi.org/10.1016/0959-4752(94)90003-5

Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-learner interaction: A decade of research. *Educational Psychology Review*, *22*(3), 271-296. https://doi.org/10.1007/s10648-010-9127-6

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Walonoski, J., & Heffernan, N. (2006). Detection and analysis of off-task gaming behavior in intelligent tutoring systems. In M. Ikeda, K. D. Ashley, & T.-W. Chan (Eds.), *LNCS, 8th International Conference of Intelligent Tutoring Systems* (Vol. 4053, pp. 382-391). Berlin, Germany: Springer. https://doi.org/10.1007/11774303_38

Wiliam, D. (2007). Keeping learning on track: Classroom assessment and the regulation of learning. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1053-1098). Charlotte, NC: Information Age Publishing.

Wilkinson, S.S. (1980). The relationship of teacher praise and learner achievement: A meta-analysis of selected research (Part of Doctoral Dissertation, University of Florida). Retrieved from: https://archive.org/stream/relationshipofte00wilk/relationshipofte00wilk_djvu.txt

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89-100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

BIOGRAPHY

Betty Tärning holds a Ph.D. in Cognitive Science and is part of the Educational Technology Group at Lund University, Sweden. Her research field covers design of virtual agents in pedagogical educational games. Of particular interest are effects of feedback a virtual agent provides.