

RESEARCH

Open Access



Games as a measure of reading and writing generalization after computerized teaching of reading skills

Ana Carolina Sella^{1*}, Jacqueline Pimentel Tenório², Carmen Silvia Motta Bandini^{3,4} and Heloísa Helena Motta Bandini⁴

Abstract

Behavior Analysis is usually accused of not being able to account for the generalization of verbal behavior that is present in linguistically competent individuals. However, several behavior analytic studies investigate this theme, and gamification has been seen as a useful way to study generalization. The purpose of this study was to evaluate reading and writing generalization in games, after these behaviors were taught through the program *Learning to Read in Small Steps*. Participants were four children between 7 and 12 years old who had reading and writing deficits. The experimental design was a pre-posttest design that encompassed five phases. Performance in probes suggests generalization of reading and writing skills to new activities (games) and responses. This study represents a small step in a systematic understanding of how games can be used to assess behavior change.

Keywords: Generalization, Reading, Writing, Children

Background

Behavior Analysis has been accused of not being able to account for the generalization of verbal behavior that is present in linguistically competent individuals (Alessi, 1987; Mackay & Fields, 2009). However, behavior analysts have come a long way demonstrating that behavioral principles and procedures, when well designed, can produce a large variety of behaviors not directly taught. These demonstrations include, but are not limited to, studies of grammar (e.g., Chase et al. 2008), vocabulary (e.g., Miguel et al. 2005), and reading and writing (e.g., Melchiori et al. 2000; Reis et al. 2009).

Generalization can be defined in several ways, and it can refer to different behavioral processes and behavior change outcomes (Cooper et al., 2007). Most commonly, when speaking of generalization within a behavior analytic perspective, authors describe generalization across time (maintenance), across settings and stimulus contexts (e.g., new people, stimuli, tasks), across behaviors (response generalization), and across subjects (Drabman et al., 1979). Specifically when it comes to stimulus

context, the specific antecedent conditions that are present when the target behaviors are taught and tested are the most important features when one wants to assess generalization. If, for example, a child is taught to how to solve multiplication facts in a horizontal format, testing in a vertical format is a change in the stimulus context, and if the child is able to correctly respond to the vertical format, generalization is shown (Cooper et al., 2007, p. 619). Because people might fail to see that differences in stimulus context are present and influencing responding, they might not program for generalization and might blame the learners for not responding according to their expectations.

One form of teaching and testing for generalization within a behavior analytic frame is to use gamification, a process defined as using game elements in non-gaming settings (Deterding et al., 2011) that has become widely used as a way to engage users and improve their experiences Hamari et al. (2014). Morford et al. (2014) defined gamification from a behavior analytic perspective as “a way to engineer the real world by arranging contingencies to bring about game-playing in a context in which game-playing does not normally occur” (Morford et al., pp. 26). In developing their definition of game-playing behaviors, Morford et al. (p. 28–30) conducted a concept

* Correspondence: carolsella@yahoo.com.br

¹Universidade Federal de Alagoas (UFAL), Av. Lourival Melo Mota, s/n.

Tabuleiro dos Martins – Maceió, Alagoas, 57010-300, Brazil

Full list of author information is available at the end of the article

analysis and described six critical features of these behaviors which consist in (a) the player's behavior directly impacting the game outcome and results; (b) the player being able to define the objectives of the game; (c) the behavior of the player being constrained by the format and contingencies in effect during the game; (d) players not being able to determine the specific outcome for each individual game; (e) the conditions in effect resulting in the development of strategies and heuristics that help in evaluating the current state of the game and in changing one's behavior to alter the probability of a certain outcome; and (f) a player not being forced to begin or end a game.

Some authors have argued that using game elements in non-gaming settings and tasks might decrease their aversiveness (de Rose & Gil, 2003; Procee et al. 2013). Even though this assertion is true in many cases, it is important to highlight the issue raised by Morford et al. (2014) that only if participants have a history of reinforcement in game-like contexts, will programming activities that recruit game-playing behaviors result in a less aversive context Bandini et al. (2006). Decreasing the aversiveness of non-gaming settings might be of particular importance in academic skill teaching contexts, especially when learners have a history of failure in performing to criterion in these contexts (Milgram et al. 1995; Solomon & Rothblum, 1984).

Specifically regarding the use of gamification to teach and test reading or writing skills within behavior analysis, there are a few examples of successful initiatives. Headsprout Early Reading® and Headsprout Reading Comprehension® are two online programs that teach learners early reading and reading comprehension skills and strategies. Learners go through different episodes that involve several game-like activities. Moving through the episodes require certain mastery criteria and is automatically mapped in a board game like screen that allows the learner to monitor their own progress. Another example is Sudo et al. (2008)'s study, which evaluated if children's spelling errors would decrease when three games (AbraKedra, memory game, and writing with chalk on the black board) involving written words, pictures, and constructed response matching to sample (CRMST) were presented. Performance in spelling tasks improved for all participants. One last example is Souza and Hubner (2010)'s study, which further evaluated the effects of the AbraKedra game on kindergartners that did not have reading or writing repertoires. The game resulted in performance improvements in both reading and writing skills. These examples, like others, taught some conditional discriminations directly and programmed learning opportunities that resulted in generalization. Sudo et al. and Souza and Hubner have explicitly programmed stimulus equivalence and minimal unit recombination procedures aiming at emergent or generalized performance.

Stimulus equivalence and recombination are widely used behavior-analytic procedures to teach reading and writing skills in Brazil de Souza et al. (2009). This occurs due to the premise that these procedures result in the emergence of novel relations without direct training (i.e., generalized performance), thus producing economy of teaching time (Alessi, 1987; Fienup et al. 2010; Sidman & Tailby, 1982). One program that has been extensively used to teach reading skills and is based stimulus equivalence and recombination is the computerized program *Learning to Read in Small Steps*, developed by Rosa Filho et al. (1998) The program teaches simple Portuguese words reading (consonant-vowel sequences) in Module 1 and complex Portuguese words reading (sequences that involve consonant clusters) in Module 2. This program has been successfully implemented with children with learning disabilities (Reis et al., 2009), preschool children (Melchiori et al., 2000), individuals with intellectual disabilities Alves et al. (2007) and illiterate adults Bandini et al. (2014) One of the main focuses of the studies using the program has been the generalization of reading and writing skills, that is, after directly teaching children how to read a few words through the program, studies have assessed generalization to new stimulus contexts (e.g., new words, in new media), to new responses (e.g., writing), and across time (maintenance). Not only studies have created their own generalization assessments, but the program itself has probes after each of the four teaching unit to assess (a) maintenance after reinforcement has been withdrawn and (b) generalization words (new words). In addition, after each teaching unit, handwriting probes are conducted to test the emergence of writing responses in a new medium (paper-and-pencil).

Even though there has been different studies evaluating the effects of the program *Learning to Read in Small Steps* Rosa Filho et al. (1998) on the generalization of reading skills across time, in a new medium (paper-and-pencil), across new words, and to new responses (i.e., writing), there has not been a study that used games to evaluate the generalization effects of the program. The purpose of this study was to evaluate reading and writing generalization in games, after these behaviors were taught through the program *Learning to Read in Small Steps*. The three games used were bingo, domino, and crosswords. The dependent variable of interest was percentage of correct responses for each response targeted in the games.

Method

Participants, setting, and materials

Participants were three girls (Dana, 9.4 years-old; Jess, 11.0 years-old; and Mona, 9.10 years-old) and one boy (Myke, 8.7 years-old), who were sent by their school

to attend a center for language studies at a university in the northeast of Brazil due to their reading and writing repertoire deficits (see Table 1). They had no other physical or neurological impairments and attended either the first or the second grade of public schools, in regular classrooms. All children were low in socioeconomic status.

Participants had to score at or below 6.7 % in the reading and writing tasks of the Reading and Writing Pretests of the computerized program *Learning to Read in Small Steps* Rosa Filho et al. (1998). The study was conducted according to the terms laid out in the approval by the university's research ethics committee, protocol #1171/09.

All sessions were about one hour long and were conducted 2–4 days per week at the center, in a room with three computers, all of which were used to deliver the teaching and testing units from *Learning to Read in Small Steps*. All computers were placed on individual desks and were equipped with a mouse, a keyboard, and a headset that delivered all of the program instructions. Session materials also included handwriting probes, which are part of the program *Learning to Read in Small Steps*, and the three target games: bingo, domino, and crosswords (Fig. 1). All games were played on a round table located on a corner of the room in which the computer desks were located. In any given session, more than one participant could be exposed to the computerized teaching procedures simultaneously, using different computers. During bingo and domino games usually more than one child took part in the probe; however only one participant was assessed in a given match. A child other than the target participant only took part in a match if they had already mastered and had been evaluated in the game that was designed to assess that given teaching unit.

Dependent Variable

In all three games, the dependent variable was the percentage of correct responses. During bingo, two sets of repertoires were evaluated, depending on the participant's choice to be a player or the caller. If a participant

chose to be the caller, target responses included drawing a piece of paper contained in a non-transparent bag, and then either (a) reading the printed word and warning the players if it was picture (e.g., “cavalo [horse], figura [picture]) or (b) reading the printed word (e.g., cavalo [horse]). If participants chose to be players, responses involved placing a bean on top of the (a) picture called or (b) word read by the caller. All participants had to play in each role at least once, so if they had already been a caller in a match, the experimenter requested that they became a player in the next match. In domino matches, target responses included placing the piece with the corresponding picture or printed word by the picture or printed word that was at either extremity of the game laid out on the table. The target conditional relations in this game were handwritten word (HW) - handwritten word (HW), handwritten word (HW) - printed word (PW), handwritten word (HW) - picture (P), printed word (PW)- printed word (PW), printed word (PW) - handwritten word (HW), printed word (PW) - picture (P), picture (P) - handwritten word (HW), picture (P) - printed word (PW), and picture (P) - picture (P). In crosswords, the target responses involved writing the missing syllables horizontally or vertically, depending on the direction indicated by the arrow placed on the top right corner of the picture (if the syllables had to be written horizontally) or the bottom right corner of the picture (if the syllables had to be written vertically).

Design and procedures

The experimental design was a pre/posttest design. The pretests included a reading test, a handwriting test, and a crossword puzzle. Performance in the bingo and the domino games were not pretested based on the assumption that if participants were not able to read the target words in the reading pretest, they would likely not be able to read these words anywhere. The procedure encompassed five general phases (see Fig. 2): (a) Pretests (*Learning to Read in Small Steps* and crosswords); (b) Unit 1 teaching and posttests from *Learning to Read in Small Steps*, followed by its corresponding game probes; (c) Unit 2 teaching and posttests, followed by its game probes; (d) Unit 3 teaching and posttests and its game probes; and (e) Unit 4 teaching and posttests and its game probes.

Pretests

The *Learning to Read in Small Steps* pretests involved asking participants to read and handwrite 15 Portuguese words that contained simple syllables (consonant-vowel syllables). Stimuli presented in this phase were those already pre-programmed in the reading software. During the reading pretest, the participant read each

Table 1 Participants' ages and scores in the reading, handwriting, and crossword pretests

| Participant | Age (years.months) | Scholar grade (years) | Pretests (% correct) | | |
|-------------|-----------------------|-----------------------------|----------------------|-------------|-----------|
| | | | Reading | Handwriting | Crossword |
| Dana | 9.4 | 1 st | 0 | 0 | 0 |
| Jess | 11.0 | 2 nd | 0 | 0 | 0 |
| Mona | 9.1 | 1 st | 0 | 0 | 0 |
| Myke | 8.7 | 2 nd | 6.7 | 0 | 0 |

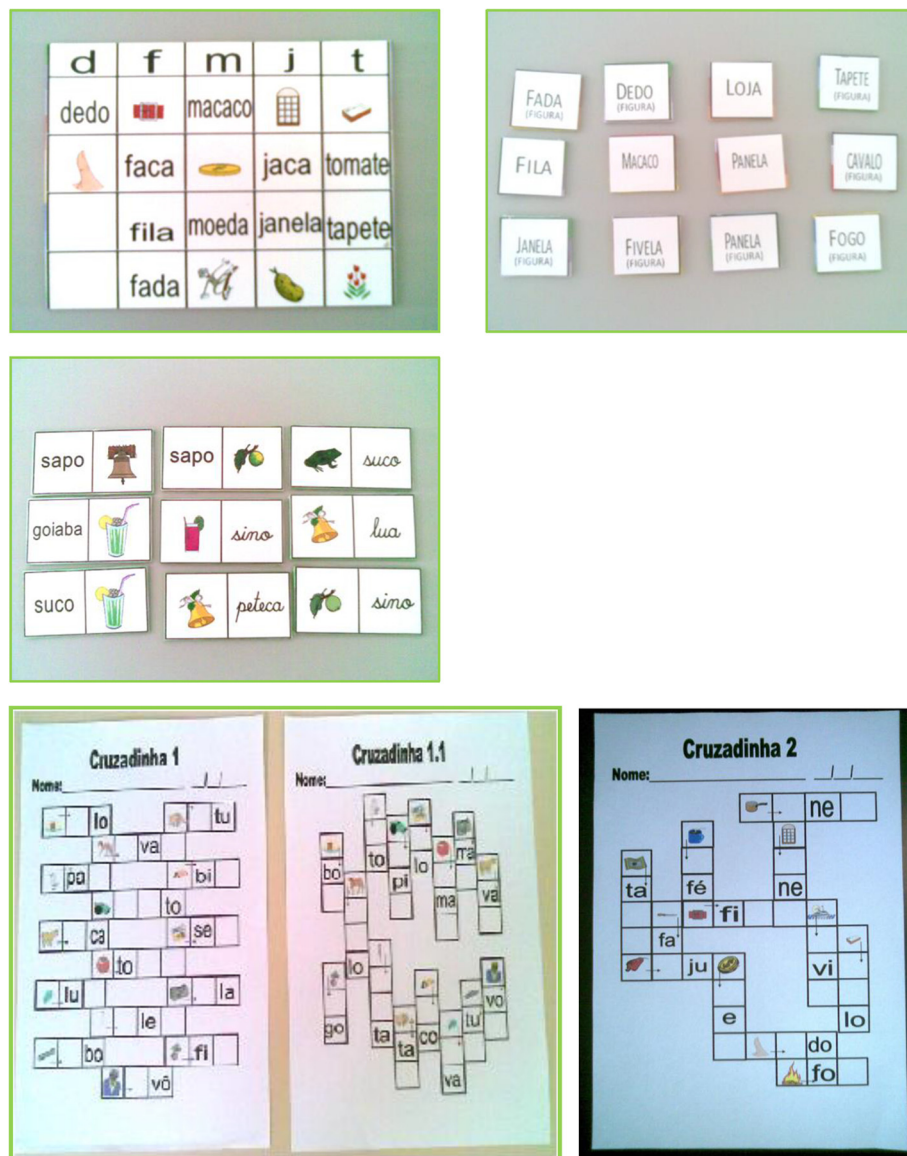
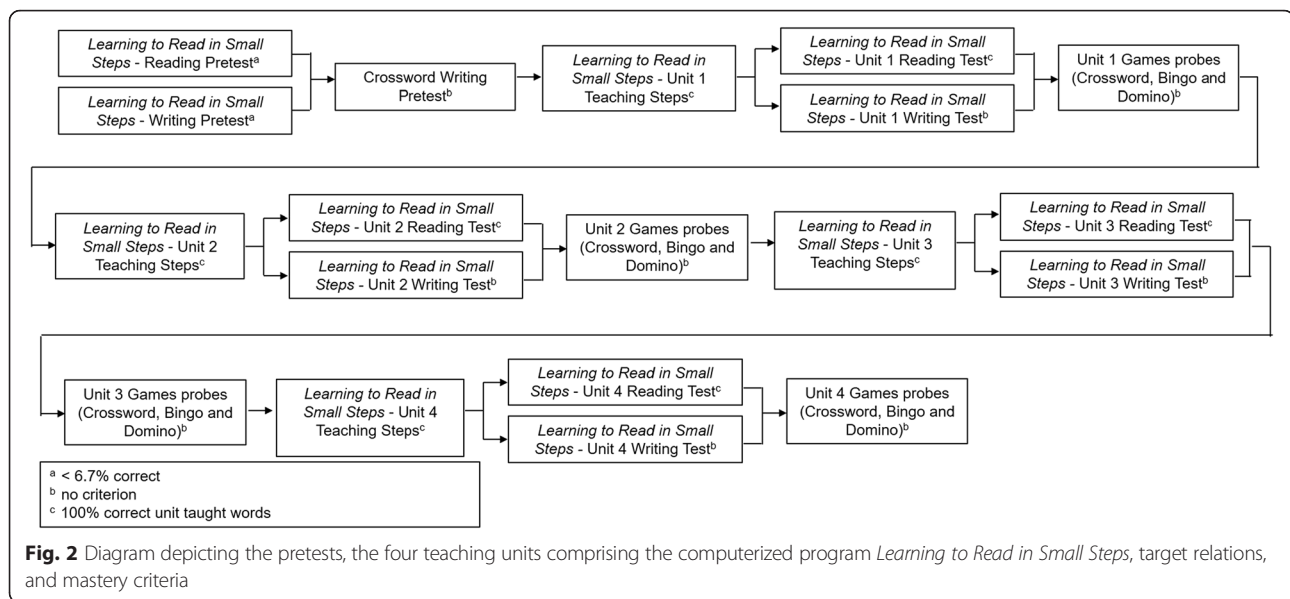


Fig. 1 Examples of each one of the games (bingo, domino, crossword)

word presented at the center of the computer screen aloud. If a participant emitted more than one correct response (i.e., 6.7 %) in the reading pretest, they were considered ineligible to participate. During the handwriting test, the participant was asked to write, in a blank sheet of paper, words that were dictated through the headset attached to the computer. There were no programmed consequences for correct or incorrect responses in this phase, just the presentation of the next trial. The crosswords pretests consisted in presenting a horizontal crossword game to the participant and asking them to complete the game. No further instructions were provided.

Program teaching units and respective posttests

The reading program procedures have been described in details elsewhere (Bandini et al., 2014). For the purpose of this study we provide a brief description of the program to establish the context in which the game probes were conducted. Module 1 of the reading program teaches 51 two-, three- or four- syllable words through steps organized into four units. Each teaching step establishes conditional relations between dictated words and printed words, dictated syllables and printed syllables, and dictated words and syllables (CRMTS). During teaching steps and their respective posttests, correct responses resulted in praise or sounds (e.g.,



hands clapping, bells tolling, etc.), respectively. Incorrect responses resulted in the computer presenting the spoken words “no, this is not it”, followed by an error correction procedure (trial repetition). Once mastery criterion was reached in the teaching steps of a unit, posttests were conducted to assess (a) percentage of correct responses in reading the words that were directly taught in the unit without differential consequences for correct and incorrect responses, (b) percentage of correct responses in reading generalization words (i.e., words that had not been directly taught), and (c) if writing responses would emerge as a result of the procedures used to teach reading skills. In this study, if a participant scored 100 % correct responses in the reading posttest, they were exposed to the game probes, and then to the next teaching unit (see Fig. 2 for the different steps of the procedure). If participants’ scored less than 100 % correct responses in a reading posttest, all words that were read incorrectly were re-taught. There was no performance criterion in handwriting posttests.

Games

Games were developed to evaluate generalization of reading and writing skills across new context, media, and responses in the three games. Game design was based on David (2001, unpublished work). Four sets of bingo were developed (see Fig. 1 for one example), one for each teaching unit. Each set was comprised of eight different cards and 46 pieces that were placed in a non-transparent plastic bag to be drawn by the caller. The cards were 10.2 cm by 13 cm and contained 25 squares with pictures and printed words (directly taught and generalization words). The pieces were 2 cm by 2.5 cm and contained the same words and pictures that were

used in the cards. As described in the dependent variable section, two repertoires were assessed in bingo probes. When participants were the caller, if they read a word incorrectly, the experimenter pointed to each syllable of the word and said “Let’s go slowly, in small steps. Do you remember this syllable?” (the experimenter pointed to the first syllable). Now, how about this syllable? (the experimenter pointed to the next syllable and repeated this procedure until the last syllable was read by the participant).“ If after this prompting procedure the participant still could not read the word, the experimenter read it to allow for the game continuation. When the participant was a player, if he or she did not place the bean on the correct word or picture, the experimenter removed the bean and placed it back at the center of the table while saying “This is not correct.” The participant was allowed to try again, even though the word was not repeated. Correct responses did not result in differential consequences and the game was over when one of the players completed the card.

Four sets of domino were developed, one for each teaching unit. Each set was comprised of 28, 4.8 cm x 9.0 cm pieces and contained only directly taught words. The pieces were divided in two parts that could contain a printed word, a handwritten word, or a picture (see Fig. 1). The game started with all 28 pieces turned upside down. Players chose seven pieces each and turned them up, so the experimenter could follow all moves. The rest of the pieces remained upside down and were used according to standard rules of domino games. Then, the experimenter turned one of the remaining pieces as the start piece. If responses were incorrect, the experimenter removed the piece, returned it to the participant that made the move and gave him or her a

second opportunity to play. Correct responses did not result in differential consequences. The game was over when one of the players did not have any pieces left or if there were no more pieces that could “be bought” to allow continuation of the game. The winner was the player that either used all of his or her pieces or the player with less pieces, if the game ended due to the lack of additional pieces.

Six crossword puzzles were developed, one containing words from all four teaching units, used as pretest, two containing words from Unit 1 (one horizontal and one vertical – these data are presented as average of correct responses in Fig. 4), and three containing horizontal and vertical words, one for each of the remaining teaching units. All puzzles were 21 cm by 29.7 cm, printed in A4-size white paper sheets and containing pictures and one printed syllable for each one of the words corresponding to the pictures (see Fig. 1). The number of syllables in each word varied from two to four. The direction in which the word should be written was indicated by an arrow placed in the same square as the picture. Before the start of each game, the experimenter instructed the participant that he or she should write the name of the picture in the direction pointed by the arrow, placing one syllable in each square. The experimenter did not provide prompts regarding which syllables were missing. If responses were incorrect, no feedback was provided since incorrect responses did not affect game continuation. Correct responses did not result in differential consequences. The game was over when the participant finished filling out all blank syllables of the crosswords.

Interobserver Agreement Scores (IOA)

All teaching procedures were computerized and were calibrated before and during data collection. Trial by trial IOA was calculated by dividing the number of trials in which there was agreement by the total number of trials multiplied by 100. In handwriting probes IOA was 100 % in the pretest; 96.7 % (range 93.3–100 %) in unit 1; 93.7 % (range 83.3–100 %) in unit 2; 95.8 % (range 91.6–100 %) in unit 3; and 91.6 % (range 83.3–100 %) in unit 4 for Dana, Jess, Mona, and Myke, respectively. In crossword puzzles, IOA was 100 % in the pretest and was 100 % in unit 1-horizontal, 94.9 % in unit 1-vertical (range 83.3–100 %), 97.9 % in unit 2 (range 91.6–100 %), 91.6 % in unit 3 (range 83.3–100 %), and 97.9 % in unit 4 (range 91.6–100 %).

Results

Overall, performance in game probes suggests generalization of reading and writing skills to the novel stimulus contexts (i.e., games) and novel responses (i.e., writing and other conditional relations).

Figure 3 shows the percentage of correct reading responses in the *program posttest* for both directly taught (performance criterion: 100 % correct responses before moving onto the next teaching phase) and generalization words (no criterion to move onto the next teaching phase). Figure 3 also shows directly taught and generalization words in *bingo* activities when the participant was a caller. All participants were able to read correctly 90.0 % or more of all words drawn during bingo matches, thus showing generalization to the new stimulus context (i.e., the bingo game) and to the new responses (i.e., reading a word from a piece). When participants were players, their performance was above or close to 90.0 % correct responses in all posttest game probes, (exceptions were Dana, 73.9 % and Myke, 82.6 % in unit 2’s posttest), also showing generalization to a new context and responses (scanning all 25 squares from a card and placing a bean over the correct word or picture).

Table 2 shows performance in domino probes. Dana and Jess were the only two participants who had the opportunity to respond to all conditional discriminations that could be present in the domino probes. Dana emitted 100.0 % correct responses in the game probes of all units in trials involving picture-picture and printed word-picture relations and 0.0 % correct responses in trials involving handwritten and picture relations. For all other conditional relations (handwritten-handwritten, handwritten-printed, printed-printed, printed-handwritten, picture-handwritten, and picture-printed) Dana’s performance varied between 0.0 % and 100.0 % of correct responses. This performance shows that generalization occurred to the new stimulus context (i.e., the domino game); however it did not occur to all new responses. Jess emitted 100.0 % correct responses in all domino probes in picture-picture relations and 0.0 % correct responses in printed-handwritten relations. In trials involving the other conditional relations her performance varied from 0.0 to 100.0 % correct responses as it happened with Dana. Again, this performance shows that generalization occurred to the new stimulus context, but not to all new responses. Mona emitted 100.0 % correct responses in all conditional discriminations to which she had an opportunity to respond, showing generalization to the new stimulus context and to all new responses that the game required her to emit. Myke’s performance was similar to Mona’s; however, in unit 1’s posttest his percentage of correct responses was a little lower: 66.7 % in handwritten word-picture relations. His performance also shows generalization to the new context and to all new responses.

Figure 4 shows two sets of data, so that the reader can compare the participants’ performance in (a) the generalization posttests that are embedded within the program *Learning to Read in Small Steps* to their

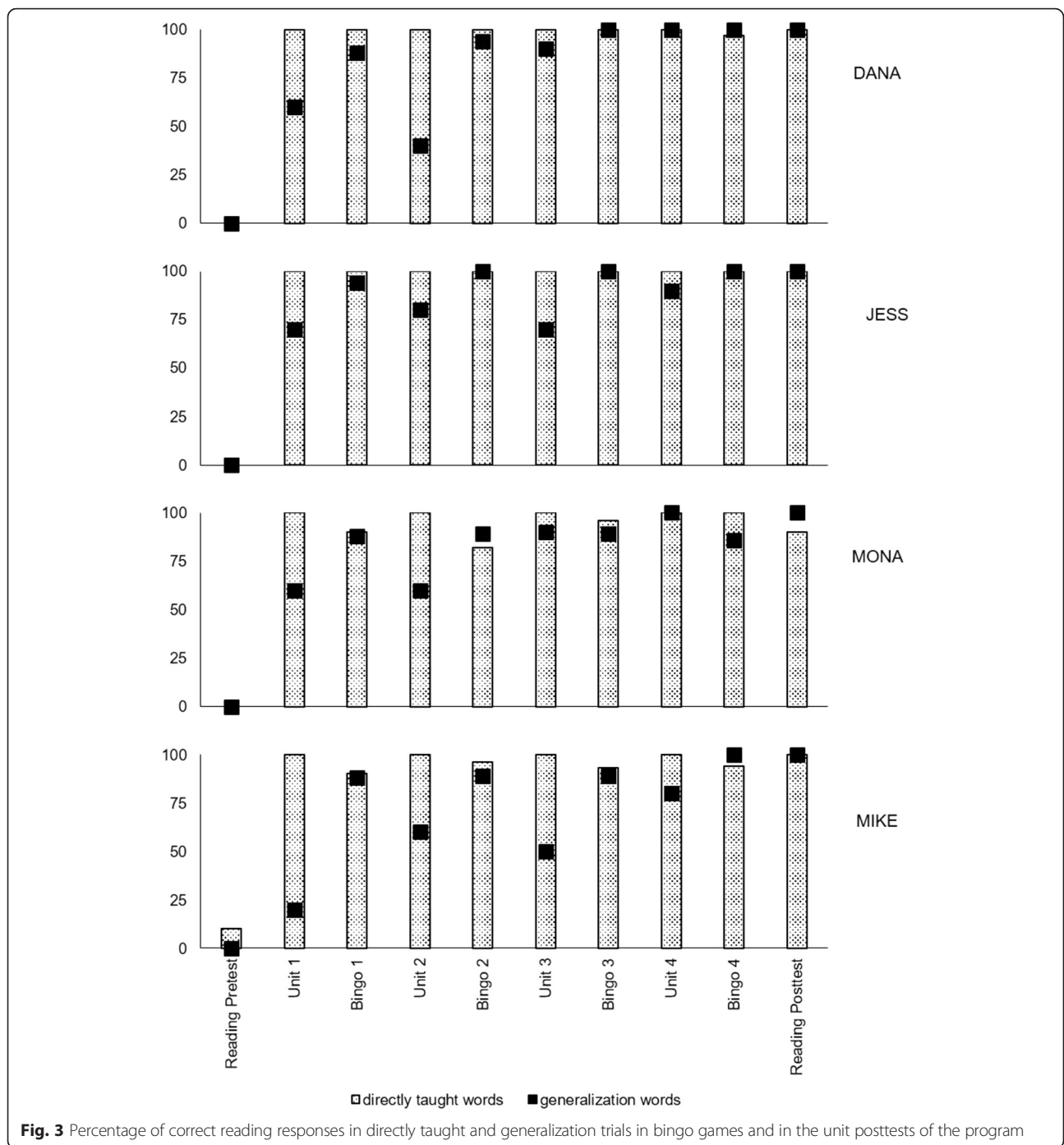


Fig. 3 Percentage of correct reading responses in directly taught and generalization trials in bingo games and in the unit posttests of the program

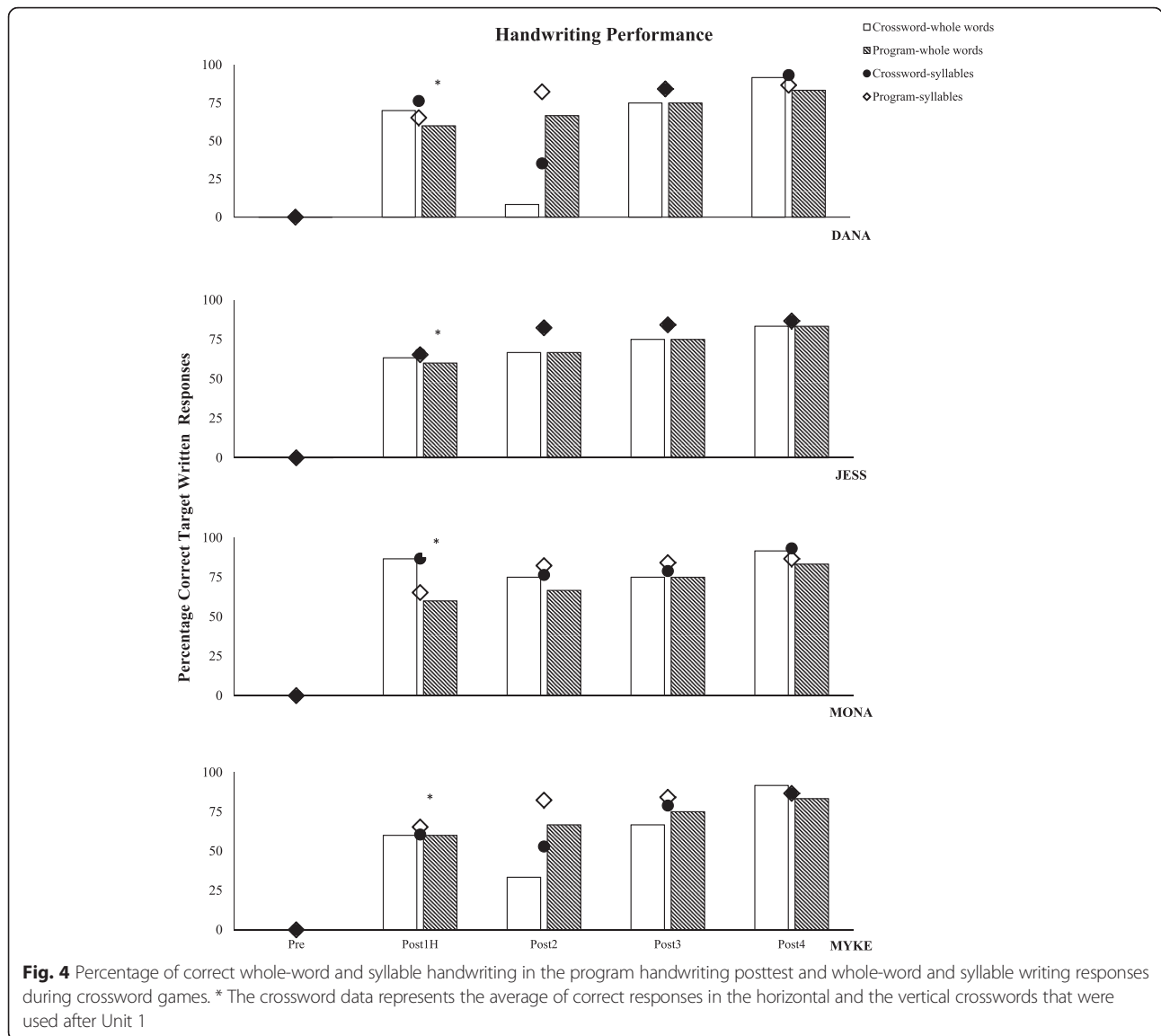
performance in (b) the crossword probes. The results are shown as the percentage of correct responses in syllables and whole words handwriting. It is important to note that even though the *Learning to Read in Small Steps* program was computerized, its handwriting posttests were conducted using paper-and-pencil. All participants performed at or above 60.0 % correct response in the first program posttest and performance improved in each new posttest, showing generalization to a new

media (i.e., paper-and-pencil) and to new responses (i.e., writing). In crossword probes, participants performed above 50.0 % correct responses, showing generalization to the new context (i.e., the crosswords game) and to the new responses (handwriting vertically and horizontally within pre-defined squares when shown a picture of the target word and an arrow to follow). Lower performances were those of Dana and Myke whose performance in unit 2's posttests was 8.5 % in whole words and 35.3 % in

Table 2 Participant scores in each conditional relation tested in domino posttest games

| Relations (abbreviation) | Domino (percentage of correct responses in posttests) | | | |
|---|---|--------|--------|--------|
| | Unit 1 | Unit 2 | Unit 3 | Unit 4 |
| DANA | | | | |
| handwritten word (HW) - handwritten word (HW) | 100 | 0 | 100 | 100 |
| handwritten word (HW) - printed word (PW) | 0 | 0 | 0 | 0 |
| handwritten word (HW) - picture (P) | 100 | 0 | 100 | 0 |
| printed word (PW) - printed word (PW) | 0 | 0 | 100 | 0 |
| printed word (PW) - handwritten word (HW) | 100 | 100 | - | 100 |
| printed word (PW) - picture (P) | 100 | 100 | 100 | 100 |
| picture (P) - handwritten word (HW) | 0 | 0 | 100 | 100 |
| picture (P) - printed word (PW) | 100 | 100 | 0 | 0 |
| picture (P) - picture (P) | 100 | 100 | 100 | 100 |
| JESS | | | | |
| handwritten word (HW) - handwritten word (HW) | 0 | 0 | 100 | 100 |
| handwritten word (HW) - printed word (PW) | 0 | 100 | 0 | 100 |
| handwritten word (HW) - picture (P) | 100 | 0 | 100 | 100 |
| printed word (PW) - printed word (PW) | 100 | 0 | 0 | 0 |
| printed word (PW) - handwritten word (HW) | 0 | 0 | 0 | 0 |
| printed word (PW) - picture (P) | 100 | 0 | 100 | 100 |
| picture (P) - handwritten word (HW) | 75 | 100 | 100 | 100 |
| picture (P) - printed word (PW) | 0 | 100 | 0 | 100 |
| picture (P) - picture (P) | 100 | 100 | 100 | 100 |
| MONA | | | | |
| handwritten word (HW) - handwritten word (HW) | 100 | - | 100 | - |
| handwritten word (HW) - printed word (PW) | 100 | - | - | - |
| handwritten word (HW) - picture (P) | 100 | 100 | 100 | 100 |
| printed word (PW) - printed word (PW) | 100 | - | 100 | - |
| printed word (PW) - handwritten word (HW) | - | - | - | - |
| printed word (PW) - picture (P) | - | - | - | 100 |
| picture (P) - handwritten word (HW) | - | 100 | 100 | - |
| picture (P) - printed word (PW) | 100 | 100 | - | - |
| picture (P) - picture (P) | 100 | 100 | 100 | 100 |
| MIKE | | | | |
| handwritten word (HW) - handwritten word (HW) | 100 | - | 100 | - |
| handwritten word (HW) - printed word (PW) | - | - | - | - |
| handwritten word (HW) - picture (P) | 100 | 100 | 100 | - |
| printed word (PW) - printed word (PW) | - | - | - | - |
| printed word (PW) - handwritten word (HW) | 100 | 100 | - | - |
| printed word (PW) - picture (P) | 67 | 100 | 100 | 100 |
| picture (P) - handwritten word (HW) | 100 | - | - | 100 |
| picture (P) - printed word (PW) | - | - | - | - |
| picture (P) - picture (P) | 100 | 100 | 100 | 100 |

The dash (-) represents data that were not obtained because in a given match there was no opportunity to respond to that specific relation



syllables, and 33.3 % in whole words and 52.9 % in syllables, respectively. Overall, generalization to new responses, media, and context was seen both in the posttests embedded within the computerized program and in game probes.

In summary, the results suggest generalization of reading skills, both in the program posttests and in new stimulus contexts (i.e., games). This can be inferred by the high scores in tasks involving directly taught words without reinforcement throughout unit posttests and game probes. The same can be verified regarding generalization words: there was a high percentage of correct responding both in the program posttests and in the bingo probes when generalization (i.e., new) words were used. In addition, writing responses, which were not directly taught by the program, were emitted in unit posttests and crossword

games, showing that by learning how to read, children could also write in different contexts.

Discussion and Conclusions

This study replicated and extended former studies by evaluating the generalization of reading and writing skills, taught through the use of the computerized program *Learning to Read in Small Steps*, across bingo, domino, and crossword games. Participants were taught to read 51 words across four teaching units that established relations among printed words, syllables, letters, pictures, and spoken words. Participants performed at 6.7 % or less correct responses in the reading, writing and crossword pretests. Then, after meeting criteria for all phases of the reading program, the data show that there was reading generalization across all three games,

including new media and words. Data also shows that writing responses also emerged, both in the program posttests and in the crossword probes. By demonstrating that the reading skills mastered within the program generalized to new contexts (games) and new responses (writing and new matching relations) this study provides additional evidence that the procedures contained in the *Learning to Read in Small Steps* program result in behavioral changes that spread beyond the program itself. This can be explained by the use of stimulus equivalence procedures in conjunction with the establishment of minimal response repertoires (e.g., de Souza et al., 2009; Hübner et al. 2009).

The games were also designed aiming to increase the likelihood that the skills that were directly taught and generalized to the game context could be also generalized to the school environment: the games were made of materials found in school settings, such as paper, EVA sheets, and contact paper. In bingo and domino probes, there were other kids playing and there was a supervising, and measures were taken to have the game occur as it would when played by fluent players (the experimenter intervened when participants' errors would influence the score or outcome of a given match). Future studies should directly evaluate generalization to game-like activities in a school context.

Specifically regarding the emergence of writing responses, as discussed in the literature (e.g., Bandini et al., 2014), CRMTS tasks and syllable training contain some correspondence with handwriting responses (word construction tasks), thus possibly making it likely that at least some generalization will occur. Another fact regarding handwriting tasks refers to the non-correspondence of performance in the handwriting program posttest versus crossword probes. While there was an increasing trend in performance in handwriting probes for all participants across units, data shows variability in performance in crossword probes.

When thinking about the critical features of game-playing behaviors as described by Morford et al. (2014), the game probes used in this study evoked behavior that met all but one criterion (non-coerciveness). Nonetheless, the authors had discussed that not all games might meet all features described in their concept analysis. During all game probes the behavior of the players had a direct effect on the outcomes and results of the game. All games had clear goals and end conditions established prior to the beginning and these were repeated at the start of each new match. All games had rules and barriers (e.g., bingo probes had a predetermined number of cards and squares that could be marked and only one piece could be drawn at a time). Bingo and domino were "more" probabilistic than the crossword probes since the final outcome could not be anticipated. For crosswords,

if participants wrote the corresponding words when presented with the specific pictures, the outcome was certain (probability of correct responses varied only due to changes in the target words). All game probes provided opportunities for development of strategies, such as looking at all pieces during domino games, counting the number of available squares in the crossword probes, and scanning the whole card during bingo. Given the possible benefits of using these games to the generalization of reading and writing skills, it is important to emphasize that, different from some of the online and board games available for purchase, these games can be designed by anyone who can follow the design instructions described in the methods section.

As highlighted in the introduction, authors have discussed the use of games as a way to decrease the aversiveness of ordinary academic tasks (de Rose & Gil, 2003; Procee et al., 2013). Again, it is important to remember that only if participants have a history of reinforcement in game like contexts, will programming activities that recruit game-playing result in a less aversive context, thus decreasing possible collateral emotional and avoidance responses that might occur during teaching, especially for children who have a history of punishment in academic contexts (Bandini et al., 2006).

This study has a few limitations that would be important to be addressed. First, the posttests (reading, handwriting, and games) were specific to each teaching unit, thus not providing a complete picture of the change occurring in participants' target repertoires. The decision to probe only words that were specific to a given unit was based on Cuvo (1979)'s and Cooper et al. (2007)'s discussion one needs to be careful when administering repeated measures for tasks that participants are not able to perform because the task might become aversive. Due to the amount of words that were involved in the probes and to the fact that all participants were receiving services because of their poor performance in academic skills, it was decided not to expose participants to probes that were likely going to result in errors. Probing all words after each teaching unit would have provided a better measure of participants' performance as they went through the steps of the reading program; however such probes might have occasioned avoidance behaviors such as disruptions or asking to leave. Future studies should attempt to monitor performance in all words after each teaching unit and, at the same time, monitor possible emotional, aggressive, or disruptive behaviors that might be incompatible with emitting the target behaviors. Future studies might also attempt to measure if by probing all words versus specific words would result in different performance due to practice effects. Another limitation has been mentioned above: to guarantee the "flow" and "fairness", the experimenter intervened whenever a participant emitted an incorrect

response. During bingo, words read incorrectly by the caller were correctly announced by the experimenter to guarantee that other players would not have an incorrect response and if players marked a wrong word or picture, the experimenter removed the bean. During domino, the wrong piece was placed back to avoid the participant winning the game with wrong combinations. Future studies should evaluate response and stimulus generalization to game probes without experimenter interference. Target responses could be monitored regardless of their effects on other people's responding or game outcomes, thus allowing for a better evaluation of the direct effects of the reading program on participant's performance in generalization probes.

Another issue that should be addressed by future studies is verifying if game-like activities are preferred over non-game like activities in the context of learning academic skills. Systematic preference assessments and concurrent arrangements in which only the activity format (game-like versus not) is changed could be helpful in addressing this question.

This study addresses Morford et al. (2014)'s suggestion of assessing generalization of behavior change through games; yet this study is only a small step in this type of investigation. As highlighted by Morford et al., there is a great need for studies that systematically investigate elements of games and how they affect behavior. Behavior analysts should undertake the challenge of designing games and studies to further our understanding on how gamification can bring about better ways to teach and maintain behaviors.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

ACS participated in the design of the study, in the data analysis and drafted the manuscript. JPT carried out the data collection, participated in the data analysis, and helped editing the manuscript. CSMB participated in the design of the study, in the data analysis, organized and supervised the data collection process, and helped editing the manuscript. HHMB participated in the design of the study, organized and supervised the data collection process, helped in the data analysis and in editing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

CNPq/FAPEAL (#2007.09.024) Ms. Carmen Silvia Motta Bandini. CNPq/INCT-ECCE (#573972/2008-7) Ms. Heloísa Helena Motta Bandini. FAPESP/INCT-ECCE (#2008/57705-8) Ms. Carmen Silvia Motta Bandini.

Author details

¹Universidade Federal de Alagoas (UFAL), Av. Lourival Melo Mota, s/n. Tabuleiro dos Martins – Maceió, Alagoas, 57010-300, Brazil. ²Universidade Federal de São Carlos (UFSCar), Rodovia Washington Luís, Km 235 - SP 310 - Jardim Guanabara, São Carlos, São Paulo, 13565-905, Brazil. ³CESMAC. Rua Cônego Machado, 918, Farol, Maceió, Alagoas, 57051-160, Brazil. ⁴Universidade Estadual de Ciências da Saúde de Alagoas (UNCISAL), Rua Doutor Jorge de Lima, 113. Trapiche da Barra - Maceió, Alagoas, 57010-300, Brazil.

Received: 9 March 2016 Accepted: 7 April 2016

Published online: 20 April 2016

References

- Alessi G. Generative strategies and teaching for generalization. *Analysis of Verbal Behavior*. 1987;5(1):15–27.
- Alves KRS, Kato OM, Assis GJA, Maranhão CMA. Leitura recombinativa em pessoas com necessidades educacionais especiais: Análise do controle parcial pelas sílabas [Recombinational reading in persons with special educational needs: Analysis of restricted syllables control]. *Psicologia: Teoria e Pesquisa*. 2007;23(4):387–98.
- Bandini CSM, Sella AC, de Souza DG. Considerações acerca do planejamento de procedimentos de ensino de discriminações complexas [Considerations regarding complex discrimination teaching procedures]. In: Guilhardi HJ, de Aguirre NC, editors. *Sobre comportamento e cognição: Expondo a variabilidade* (vol. 17). Santo André, SP: Esetec; 2006. p. 61–71.
- Bandini CSM, Bandini HHM, Sella AC, de Souza DG. Emergência de leitura e escrita em adultos analfabetos após tarefas de matching-to-sample [Emergence of reading and writing in illiterate adults after matching-to-sample tasks]. *Paidéia*. 2014;24(57):75–84.
- Chase PN, Ellenwood DW, Madden G. A Behavior Analytic Analogue of Learning to Use Synonyms, Syntax, and Parts of Speech. *The Analysis of Verbal Behavior*. 2008;24(1):31–54.
- Cooper JO, Heron TE, Heward WL. *Applied behavior analysis*. 2nd ed. Upper Saddle River, NJ: Pearson; 2007.
- Cuvo AJ. Multiple baseline designs in instructional research: Pitfalls of measurement and procedural advantages. *Am J Ment Defic*. 1979;84(3):191–203.
- de Rose JCC, Gil MSCA. Para uma análise do brincar e de sua função educacional [For an analysis of playing and its educational function]. In: Brandão MZS, editor. *Sobre comportamento e cognição: A história e os avanços, a seleção por consequências em ação* (vol. 11). Santo André, SP: ESETEC; 2003. p. 373–82.
- de Souza DG, de Rose JC, Faleiros TC, Bortoloti R, Hanna ES, McIlvane WJ. Teaching generative reading via recombination of minimal textual units: A legacy of verbal behavior to children in Brazil. *Int J Psychol Psychol Ther*. 2009;9(1):19–44.
- Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D. Gamification: Using game-design elements in non-gaming contexts. In: CHI'11 Extended Abstracts on Human Factors in Computing Systems; New York, NY: ACM. 2011. p. 2425–8.
- Drabman RS, Hammer D, Rosenbaum MS. Assessing generalization in behavior modification in children: The generalization map. *Behavioral Assessment*. 1979;1:203–19.
- Fienup DM, Covey DP, Critchfield TS. Teaching brain—behavior relations economically with stimulus equivalence technology. *J Appl Behav Anal*. 2010;43(1):19–33.
- Hamari J, Koivisto J, Sarsa H. Does gamification work? A literature review of empirical studies on gamification. USA: Proceedings of the 47th Hawaii International Conference; 2014. p. 3025–34.
- Hübner MMC, Gomes RC, McIlvane WJ. Recombinative generalization in minimal verbal unit-based reading instruction for pre-reading children. *Experimental Analysis of Human Behavior Bulletin*. 2009;27(1):11–7.
- Mackay HA, Fields L. Syntax, grammatical transformation, and productivity: A synthesis of stimulus sequences, equivalence classes and contextual control. In: Rehfeldt RA, Barnes-Holmes Y, editors. *Derived relational responding: Applications for learners with autism and other developmental disabilities: A progressive guide for change*. Oakland, CA: New Harbinger; 2009. p. 209–35.
- Melchiori LE, de Souza DG, de Rose JC. Reading, equivalence, and recombination of units: A replication with students with different learning histories. *J Appl Behav Anal*. 2000;33(1):97–100. doi:10.1901/jaba.2000.33-97.
- Miguel CF, Petursdottir AI, Carr JE. The effects of multiple-tact and receptive-discrimination training on the acquisition of intraverbal behavior. *The Analysis of Verbal Behavior*. 2005;21:27–41.
- Milgram N, Marshevsky S, Sadeh C. Correlates of academic procrastination: Discomfort, task aversiveness, and task capability. *J Psychol*. 1995;129(2):145–55.
- Morford ZH, Witts BN, Killingsworth KJ, Alavastous MP. Gamification: The intersection between behavior analysis and game design technology. *The Behavior Analyst*. 2014;37(1):25–40.
- Procee R, Kamphorst B, Wissen AV, Meyer JJ. A formal model of procrastination. Proceedings of the 25th Belgium-Netherlands Artificial Intelligence Conference; 2013. p. 152–9.
- Reis TS, de Souza DG, de Rose JC. Avaliação de um programa para o ensino de leitura e escrita [Evaluation of a program to teach reading and writing]. *Estudos em Avaliação Educacional*. 2009;20(44):425–52.

- Rosa Filho AB, de Rose JCC, de Souza DG, Hanna ES, Fonseca ML. Aprendendo a ler e a escrever em pequenos passos [Learning to read in small steps - computer software]. 1998.
- Sidman M, Tailby W. Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *J Exp Anal Behav.* 1982;37(1):5–22.
- Solomon LJ, Rothblum ED. Academic procrastination: Frequency and cognitive-behavioral correlates. *J Couns Psychol.* 1984;31(4):503–9.
- Souza SR, Hubner M. Efeitos de um jogo de tabuleiro educativo na aquisição de leitura e escrita [Effects of an educative board game on the acquisition of reading and writing]. *Acta Comportamental.* 2010;18(2):215–42.
- Sudo CH, Soares PG, Souza SRD, Haydu VB. Equivalência de estímulos e uso de jogos para ensinar leitura e escrita [Stimulus equivalence and using games to teach reading and writing]. *Revista Brasileira de Terapia Comportamental e Cognitiva.* 2008;10(2):223–38.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
