International Journal of Child-Computer Interaction 14 (2017) 15–22

Contents lists available at ScienceDirect

International Journal of Child-Computer Interaction

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

# Game-influenced methodology: Addressing child data attrition in language development research

Tania S. Zamuner *, Lara Kilbertus, M. Weinhold

*University of Ottawa, Canada*

## HIGHLIGHTS

- Task engagement impacts the efficacy of developmental research with preschoolers.
- Preschoolers in the gamified group outperformed the control group on task performance.
- Gamification of language-based task improves preschoolers' attention and engagement.
- Provides a beginning point for CCI using children's verbal responses as an interactive tool.

## ARTICLE INFO

*Article history:*
Received 5 November 2015
Received in revised form 3 July 2017
Accepted 1 September 2017
Available online 5 October 2017

*Keywords:*
Preschool children
Language development
Experimental research
Gamification
Gameful design
Motivation
Assessment

## ABSTRACT

Research on human development can be challenging for many reasons, one of which is high attrition rates for infants and children. To address the issues of attention and engagement, we examined whether gamification of an experimental methodology improved preschoolers' participation on a task. The Primed Picture-Naming Task (PPNT) has been widely used to study language processing in adults; however, few studies have successfully implemented the methodology with children under 5 years of age, in part due to children's difficulty in completing the task. One version of the PPNT incorporated narrative and goal-directed game-like elements, while the control version was presented in a traditional format. The results showed that preschoolers' participated longer and completed more trials compared to children in the control version. Gamification is a valuable tool for creating assessments of preschoolers' language development and for improving young children's motivation and engagement in research. This study provides a beginning point for the development of Child–Computer Interactions which use children's verbal responses as an interactive tool.



## 1. Introduction

Studying human development is challenging for many reasons, one of which is the high attrition rates for infants and children participating in experimental tasks [1–3]. While adults are able to focus their attention during long tasks, children require truncated versions. For example, experimental tasks with adults often last over 30 min, while experimental tasks for children are typically designed to last not more than 15 min. Moreover, adults and children cannot complete the same complexity of tasks, and tasks often need to be changed or simplified for younger participants (e.g., [4,5]). The difficulty of a task can also mask learners' abilities [6,7]. This makes comparisons between adults and children potentially difficult, since methodological changes can compromise the experiment's validity in testing the same phenomenon. Even if children are capable of completing a complex task, their performance can be influenced by attention and motivation [8]. One solution to address these issues is gamification, which has been defined as "the use of game design elements in non-game contexts [9, 9]". In this study, we examined whether gamification of an experimental methodology improved preschoolers' participation on a task. We used the Primed Picture-Naming Task (PPNT), which has been widely used to investigate language processing in adults. In this task, participants hear an auditory prime ('ball') followed by a target image which they name (BOOK). Researchers have had mixed success with the PPNT methodology to study language development. Moreover, few studies have successfully implemented the methodology with children under 5 years of age, in part due to children's attention and difficulty with completing the experimental task. In this study, we examine the effects of gamification on 2- to 4-year-old children's participation in the task by comparing two versions of a PPNT experiment. One version incorporated both narrative and goal-directed elements of game design, while the control version was presented in a simpler,

* Correspondence to: Department of Linguistics, University of Ottawa, Arts Hall, 70 Laurier Ave. East, Ottawa ON, Canada K1N 6N5.
*E-mail address:* tzamuner@uottawa.ca (T.S. Zamuner).

https://doi.org/10.1016/j.ijcci.2017.09.001
2212-8689/

more traditional format. The objective was to determine whether gamification of the PPNT impacted preschoolers' attention and engagement on the task. This study provides insight for Child–Computer Interactions as we explore children's performance in producing verbal responses to visual and audio computer prompts, across signal/communication modes.

### 1.1. Challenges and methodologies in developmental research

Rates of attrition and data loss in developmental research are consistently higher than in studies with adults. For example, Zamuner et al. [10] conducted the same word recognition experiment with adults and 2- to 3-year-olds. All of the adults completed the task, but 7% of the 2- to 3-year-olds did not. There is some variation on attrition and data loss with children depending on the type of methodology and age of children. For example, the word recognition study just described is an easy task to implement with young children and has relatively low attrition rates. However, more complex tasks such as the two-alternative forced-choice picture-matching task (2AFC) has much higher attrition rates with children the same age. Zamuner [11] reports that 65% of 2-year-olds run on the 2AFC were excluded from the analysis and Skoruppa, Mani and Peperkamp [12] report that 50% of 2- to 3-year-olds were excluded. A variety of reasons are provided for why the participants were excluded, such as failing to complete a pre-test, failing to complete the experimental task due to fussiness or lack of interest, and failing to complete enough experimental trials. Results can be skewed if there is an attrition bias, for example, if the youngest participants drop out of a control condition, but not a test condition, age becomes a confounding factor with the tested variable [8,13]. Skewed data requires complex statistical analyses [14] or testing more children to make up for lost data. Child participants may also complete only part of a study, and depending on the structure of the trial orders, this may cause an imbalanced number of data points in different conditions. For many methodologies, high rates of infants and children are excluded from studies due to fussiness or an inability to complete the experimental task. These high attrition rates indicate the necessity of revising designs to make them more appealing to children [1,3,13,15]. However, children's engagement cannot be entirely attributed to an inability to focus their attention or perform a repetitive task. When it is an activity that children enjoy, they can play for long periods of time, which can be extreme in some cases [16]. If experimental tasks can be made appealing, children should be more likely to complete the entire study. This would result in cleaner and more data to analyze, more efficient use of available participants, and a more positive experience for children.

In the field of language development, assessments are increasingly being administered with the use of computers and tablets [17]. This presents opportunities to improve methodological paradigms, reduce attrition rates, and improve attention and performance. Many research paradigms require controlled presentation of stimuli and timing of trials. By administering an experimental task on a computer or a tablet, experimenters can guarantee that the testing materials are the same for each child. To date, the majority of research looking at developmental speech production is longitudinal or uses off-line experimental methodologies. For example, children's speech production skills have been evaluated by looking at what words are added to children's vocabulary over time from data collected in naturalistic settings (e.g., [18,19]). Other experimental assessments of children's verbal language skills focus on global measures, such as analyzing children's responses as either correct or incorrect. These coarse-grained measures are not as sensitive compared to fine-grained, on-line or implicit measures of language processing [20]. An example of the difference between global versus fine-grained measures can be seen in the Picture Naming Task (PNT), where a child is asked to name the pictures, e.g., 'table' and 'piano'. Based on global measures, a 4-year-old will produce both words accurately. However, using fine-grained measures, there is actually a difference in how children perform on these trials; children generate or produce the early-acquired word 'table' faster than the late acquired word 'piano' [21]. The dependent measure used to calculate the speed of speech production is the speech reaction time (SRT): the mean naming latency or delay in milliseconds between the presentation of the image and the beginning of the utterance [22]. An adaptation of the PNT is the Primed Picture Naming Task (PPNT). In priming tasks, one stimulus (the prime) precedes another (the target). When the prime and target share an attribute (e.g., the same initial sound as in 'cat' and 'cow', or semantic features as in 'cat' and 'dog'), SRTs in the shared-attribute or 'related' condition are different than in the unrelated condition. The direction of the effect depends on the stimuli and the timing of the prime-target (e.g., [23,24]). In some studies, SRTs in the related condition are faster than in the unrelated condition which is called facilitative priming, and in other studies SRTs in the related condition are slower than in the unrelated condition and this is called inhibitive priming. Effects of priming in the adult literature are well established [25,26], and have provided many insights into how language is represented. For example, results from priming studies have revealed that our mental lexicon is an interconnected network with links between related words [27]. When we hear the word 'dog', our mental concept for 'dog' is activated, which in turn activates words related in sound (e.g., 'doll', 'dot', 'fog') and meaning ('cat', 'canine').

There are some published studies using SRT measures with typically developing children, children who stutter, and children with autism spectrum disorder, and results have typically reported facilitatory phonological priming effects (E.g., [28]). However, the existing handful of studies have almost exclusively been done with children aged 4 to 5 years and older [21,28–33], with no studies with children between 2 to 3 years of age. Data from these groups are important for education because the language skills acquired in the early preschool years form the basis for later literacy skills, such as reading and writing [34,35]. While the PPT and PPNT are promising methodologies for learning about children's language development, one of the central difficulties in using these methodologies is in engaging children in the task. In the literature, numerous studies have reported that preschoolers were not interested in participating in the task [21] and that attrition rates were due to "various maturational, attentional, and behavioral variables exhibited by 3–5-year-old children, [28]: 1432". Thus, the challenge of conducting the PPT and PPNT with young children makes these tasks a good choice for gamification, with the goal of improving children's attention, motivation, and engagement in the task. In addition, there are almost no experimental methodologies that are suitable to use with children under 3 years of age to examine on-line processing during speech production. The ability to implement the PNT and PPNT with this younger age group would enable researchers to examine how words are activated and processed during on-line speech processing.

### 1.2. Gamification

Gamification refers to "the use of game design elements in non-game contexts ([9]: 9)". Game-design elements are typically introduced to increase motivation and gamification has been applied to a variety of contexts, such as service and marketing, task management, and education [36–38]. See Deterding et al. [9] for a hierarchy of game design elements. Dondlinger [39] and Gee [40] describe different ways of increasing motivation and enjoyment for the player within an educational context, including the creation of an interesting story to frame the game; the inclusion of goal-focused elements such as a score, number of levels completed,

progress bar; and being able to learn the game tasks while playing a tutorial. Broadly characterized, two effect components of game design are Narrative and Goal elements [39–41]. Narrative games such as Dragon Age [42] are structured around a story, with little focus on game mechanics or scoring points, while Goal games like Angry Birds [43] constantly reward players with points, stars, and other prizes as they finish a level. Each style has distinct characteristics which differ in how easily they may be adapted to an experiment design. Goal design elements indicate the player's progress and success throughout the game, providing information about the number of levels completed and displaying a cumulative score. Malone [37] examined the success of goal-related features in keeping 11-year-old children's attention. He found that the preferred goal-design elements were those which provided feedback on player progress (e.g., scorekeeping; graphics representing progress towards the endpoint). The ability to change the outcome of the game ("interactivity") was also an important feature [37]. However, Dondlinger [39] argues against trying to increase participation by attaching a score to tasks or offering intangible trophies for completion, claiming that these attributes are successful when integrated into a complete game, but do little to increase motivation when tacked on to a task without the broader context of a game's framework. Narrative games fit the game tasks into an overarching story, situating the player as an active participant in the game world who is able to affect outcomes by interacting with the game's story and making choices. Waraich [44] incorporated narrative-driven elements to a university classroom to increase learning and attention. He did this by formatting computer science quizzes as a space-themed game. An interesting result was that the graphics quality did not impact participants' interest in the game as much as the storyline. This is a promising finding for researchers wishing to add elements of game design to their experimental paradigms because it suggests that this can be achieved with graphically simple displays.

Most gamification has been applied to increase adults' interest and motivation, though a few studies have introduced game elements to tasks for school-aged and younger children. This has been achieved by adding goal-focused elements of game design. For example, in a touchscreen-gesture task, Brewer et al. [15] dramatically increased the number of 6-year-olds who completed the full experiment (224 trials in two different tasks) when they added a 'score' number to a corner of the touchscreen display. While 50% of task sessions in the control version were completed, 92% of sessions were completed in the gamified version. Note that children could also trade in the points they earned during the experiment for a small prize at the end of the session, and the authors did not address whether the ability to win a prize might be a greater factor in increasing children's motivation than the scoring component.

### 1.3. Gamification of research on language

More specifically related to the purpose of the current study, previous research has also incorporated elements of game design into assessments and research on language abilities and language development. Achim and Marquis [45] designed an automated version of the Bilingual Aphasia Test (BAT) to be administered on a computer, which had advantages over pen-and-paper format because it was easier to score, more accessible to participants who had difficulty writing, and allowed for a quicker calculation of results. Gamification can improve participants' cooperation and make clinical assessments easier to administer, particularly when working with children. Marquis [46] found increased interest and participation when administering the BAT to school-aged children through a touch-screen tablet, and credited the improvement to the novelty of the technology and that the tablet version of the BAT was similar to existing tablet games. Similarly, Bello et al. [47] developed a computerized alternative to the Italian MacArthur–Bates Communicative Development Inventory (MB-CDI). The traditional format of the MB-CDI is a pen and paper checklist that parents/guardians fill out. In the computerized version, children from 19 to 30 months old were presented with three pictures (a target, a semantic distracter, and an unrelated image) and asked to point to the verbally labeled target. After children pointed to the target image, the target and the distracter images disappeared, and the label for the remaining unrelated image was elicited. In total, the session took 30–40 min to complete. This alternation of production and comprehension was meant to infuse variety in the task, as well as mimic the turn-taking of natural speech. A statistical comparison of children's results for the picture-naming game and their scores on the MB-CDI as completed by parents/guardians found significant correlations between the two, suggesting that the game could be a viable alternative to traditional methods of vocabulary measurement [47].

Lastly, research in language development has explored gamification as an avenue for improving participation rates and the quality of the research data collected. Polišenská and Kapalková [48] sought to improve cooperation in non-word repetition tasks with children aged 2 to 6 years of age. They included both narrative and goal elements, creating a story in which the child participant could repair a broken necklace by repeating the 'magic words' they heard. Each accurate repetition was rewarded with another bead being added to the display at the bottom of the screen, providing a clear interactive element and feedback on progress over the duration of the experiment. Their participant dropout rate was only 1.8%, compared to the average of 18.3% for similar studies, suggesting this would be a viable approach for similar experimental methodologies. Children also repeated more items on Polišenská and Kapalková's task compared to other studies. The increase in amount of data collected also improves on the generalizability of the results.

Although the studies above show improved performance with the introduction of gamified elements, adding game elements to a task may potential increase its complexity. With added complexity, one runs the risk of changing the methodology and failing to test the intended variables. For example, participants may be distracted by their enjoyment and no longer focus sufficient attention on performing the experimental task. Hawkins et al. [49] compared gamified and ordinary experiments with adults, testing whether introducing game-elements such as a score, backstory narration, and animations affected the quality of the data collected. They found no significant differences in the quality of the data collected between the gamified and ordinary experiments. No studies have examined whether gamification of an experimental task has an effect on the data collected from young children. This question is pursued in the following experiment.

## 2. Experiment

To date, elements of game design have infrequently been combined with experimental methodology, highlighting the need to compare gamified and traditional methodologies. This study compares a version of the PPNT that incorporates both narrative and goal-directed elements of game design, while the control version is presented in a simpler, more traditional format. We predict that preschoolers tested on the gamified version will complete more trials compared to preschoolers tested on the control version. This is expected based on the success of Brewer et al. [15] and Polišenská and Kapalková [48] who incorporated game elements into their experimental designs. We also investigate whether gamification affects the quality of the experimental data by examining whether there are any relationships between the number of experimental

trials completed by the children and the consistency of the data collected. Lastly, the PPNT task contains both related prime-target pairs (e.g., ball-BOOK) and unrelated prime-target pairs (e.g., cake-BOOK). Although this methodology has not been previously used with 2- to 4-year-old children, we predict that preschoolers would have different speech reaction times (SRTs) to naming images when preceded by related primes compared to unrelated prime conditions, in keeping with established priming research with older children.

### 2.1. Methods

#### 2.1.1. Participants

Participants were 16 children matched in age (±2 months, $M = 1$ month), and gender (total of 6 girls, 10 boys, 3 girls and 5 boys in each group). Children in the gamified condition had a mean age of 37.14 months (range = 30–54 months, $SD = 8.5$ months) and children in the control condition had a mean age of 37.19 months (range = 29–52 months, $SD = 7.8$ months). All participants were native speakers of English without diagnosed language or developmental delays, as reported through a parental questionnaire.

#### 2.1.2. Materials

The stimuli were 20 words (*book, ball, car, cake, dog, doll, mouse, moon, pig, peas, cheese, chair, light, lips, rock, rain, sheep, shoes*). Each word was assigned a Related prime, in which the prime and target (indicated in uppercase) matched in first consonant (ball-BOOK), and an Unrelated prime which did not match (cake-BOOK). There were 40 trials. Each word appeared once as a prime and once as a target in the first 20 trials (e.g., Related pig-PEAS, unrelated doll-PIG). In the second 20 trials, each word appeared in a different Relatedness condition (e.g., unrelated pig-MOON, related peas-PIG). Unrelated primes did not rhyme with the target, and were semantically unrelated. Five additional words were used for practice trials. There were 4 test lists to counterbalance the order in which the stimuli were presented. In previous studies with young children, priming condition is presented as a blocked condition, i.e., a block of related trials, followed by a block of unrelated trials, with or without counterbalancing across participants for which block is presented first. Note that our designed mixed the presentation of the different priming conditions. This is more similar to the designs used with adults, in which different priming conditions are presented in a randomized order.

Audio stimuli were recorded using a female English speaking adult and controlled for pitch, amplitude, and duration using Praat [50]. The 5 practice words were modified with Praat to have a high average pitch to make the words sound cartoon-like. Full-color clipart images were used for the 20 experimental words. In the control condition, a purple arrow was used to cue the participant to respond. In the gamified condition, there were three images of Pingu (The Pygos Group). There were also eight 5-second video clips of Pingu performing various activities (e.g. skiing, fishing, dancing) for inter-block screens. The experimental was presented using Experiment Builder Software (SR [51]) installed on an Apple Mac Laptop. Audio recordings of participants' responses were collected through Experiment Builder using a SHURE Microflex MX392/0 omnidirectional microphone and a Blue Microphones ICICLE XLR to USB adapter. We coded children's verbal responses off-line to bypass potential issues when using automatic voice trigger keys [52]. Automatic trigger systems can be especially problematic when used with young children since any noise (and not always the desired speech response) can trigger the voice key.

#### 2.1.3. Procedure

Children were first familiarized with the stimuli and asked to name all practice and test images, presented on PowerPoint. The control condition of the task was then demonstrated to all children using the practice stimuli, again presented on PowerPoint. After this training, participants were invited to play a video game in a sound-attenuated testing booth. In the Gamified condition, children were told they would be helping Pingu learn some words by saying what the picture was and ignoring what the computer said. In the Control condition, children were told they would name pictures while ignoring what the computer said. The first 5 trials demonstrated the task using the pitch-modulated, cartoon-sounding practice items. In the Gamified condition, an image of Pingu speaking with a speech bubble was displayed when the target image appeared, and the correct word for the target image was played after a 500 ms delay (this gave the sense that Pingu was performing the task and naming the pictures). The timing of the trials in the Control condition was the same, but the speech bubble appeared without the image of Pingu (see Fig. 1). The training trials were followed by 5 practice trials and then 40 test trials, all using the same procedure. Trials began with an animated fixation point. The experimenter triggered the start of the trial with a keypress when the child was looking at the screen, and this began the audio recording of the trial. After 200 ms the audio prime played, which lasted about 500 ms. The target image appeared 450 ms after the offset of the prime and a visual cue pointed at the image. This meant that there was no overlap between the audio prime and the target image. This timing was chosen because Melnick et al. [28]: 1431–1432) reported that in pilot testing, when the audio prime and target image overlapped, 3- to 4-year-old children had less success in completing the task. The image remained on the screen for 5000 ms or until the participant responded, at which point the experimenter triggered the next trial. In the Gamified condition, Pingu was shown on the screen while the audio prime was being played. When the target image appeared on the screen, Pingu pointed to indicate that children should name the image on the screen. In the Control condition, a purple line was shown while the audio prime played. When the target image appeared on the screen, a purple arrow pointed to the target image (Fig. 1, a generic penguin image has been used for the purpose of the figure).

In both versions, a video clip of Pingu appeared after every 5 trials, followed by an inter-level display. In the Gamified condition, the inter-level display comprised an array of ten stars, a 'Score' number, a red X labeled 'Quit', and a green arrow labeled 'Next'. The background screen was of a winter forest. After each block, one of the stars would change from gray to yellow and the 'Score' number would increase, accompanied by a 1000 ms musical clip. In the Control condition, an inter-level display appeared which contained a red X and a green arrow, and the background screen was white. In both conditions, the participants were asked if they wanted to keep playing. If they answered yes, the experimenter continued; if they answered no, the experiment was terminated.[1] At the end of the control version, the experimenter ended the study. At the end of the gamified version, the tenth star on the inter-level display turned yellow, and the text "YOU WIN!" appeared, along with an image of Pingu cheering.

[1] Note that the methodology of asking participants whether or not they want to continue is not a typical procedure when running an experiment with children. Rather, children are typically encouraged to participate for as long as possible. We thank an anonymous reviewer for raising this point. It is possible that if this component of the procedure were removed, children in the control study might complete more trials.

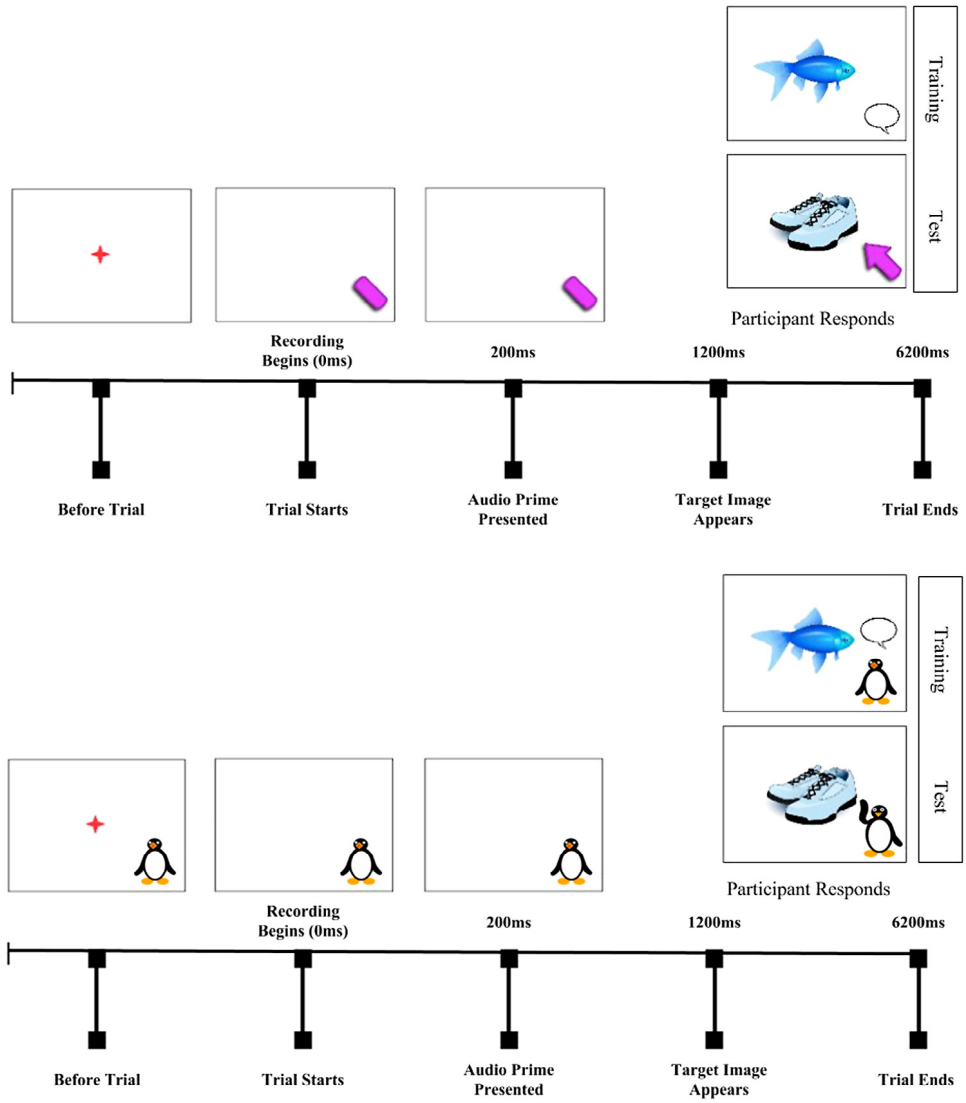


**Fig. 1.** Timeline for training and test trials for control version (top panel) and game version (bottom panel) of the PPNT.

## 3. Results

The data were first coded for the number of trials each participant completed. Each child in the gamified condition completed all 40 trials, compared to children in the control condition who completed on average 27 trials (range 15–40 trials, $SD = 11.7$). Trials were then coded for accuracy and the number of usable trials. In studies using SRTs, analyses are only done on correctly produced trials; therefore, it was important to determine whether the Game and Control versions of the experiment would impact the quantity of usable data. Trials were coded as incorrect when participants did not correctly label the target image (e.g., 'lamb' for 'sheep'), talked over the prime, did not give a response, or used a determiner phrase or a carrier phrase (e.g., 'a pig'). After data exclusion for trials that were not usable for the above reasons, children in the gamified version had significantly more usable trials ($M = 28.75$, SD $= 11.29$) than children in the control version ($M = 14.5$, SD $= 12.26$), $t(7) = 2.89$, $p < .05$, $d = 1.45$. The accuracy of children's responses was the same in both conditions. In other words, children in the gamified condition were not more accurate than children in the control condition; however, because children in the gamified condition completed more trials overall, they had more usable data. Lastly, while aged was matched across conditions, analyses looked at whether there was a correlation between age and the number of usable trials. In the gamified condition, there was no correlation between the participants' age and the number of usable trials ($r(6) = .25$, $p > .05$); however, in the control condition, there was a weak positive correlation indicating that older children contributed more usable trials than younger children ($r(6) = .62$, $p = .1$).

One concern in adding game elements to an experiment is that the quality of the observations might be influenced by the experimental task. Although individual participants' rate of accuracy was not affected by the version they were tested on, it is possible that since the length of participation varied between the conditions, there might be differences in average SRTs between the groups. Participants who participated longer and completed more trials might show a fatigue effect, which could result in slower responses towards the end of the experiment. Another possibility is that participants' responses could become faster due to practice effects. To evaluate whether the length of time spent performing the experiment was correlated to children's speed of responses, a comparison was made between trial number and children's SRTs. As a reminder, SRT refers to the amount of time in milliseconds between the presentation of the image and the beginning of the child's speech response, i.e., the amount of time it takes children

to name the image. There was no correlation between trial number (e.g., trial 1, trial 2, etc.) and SRTs in the gamified condition ($r(221) = .004$, $p > .05$) and no correlation in the control condition ($r\ (114) = -.04$, $p > .05$). This means that participants were not slower or faster at responding over the course of the experiment. There was also no difference in the average speed of children's SRTs in the gamified condition ($M = 1122$ ms, $SD = 416$ ms) compared to the control condition ($M = 1088$ ms, $SD = 593$ ms).

Having examined how accuracy and length of participation were affected by the between-subject variable of game versus control condition, we turn now to the within-subject variable of priming condition (related, unrelated), which was predicted to have an effect on children's SRTs. Table 1 provides a summary of the SRT analyses. Recall that SRT analyses are only conducted on correct responses and children in the gamified condition produced more usable trials. Following Melnick et al. [28], participants were excluded from the analysis if they had less than 33% usable data (also see [29,32]. Based on this criterion, 1 participant (12.5%) was removed from the gamified condition and 5 participants (62.5%) were removed from the control condition. The remaining participants in the gamified condition were on average 7 months younger than those in the control condition. In the gamified condition, there was a significant effect of priming ($t\ (6) = 2.62$, $p < .05$, $d = 0.49$), with 5 of the 7 children patterning in the same direction, meaning that these individual children had higher SRT means in the related versus unrelated priming conditions. Children were on average 117 ms slower to name the image of the 'BOOK', if they were primed with a word beginning with the same sound ('ball') compared to trials on which they were primed with an unrelated sounding prime ('cake'). While the data patterned in the same direction with participants in the control condition, with slower SRTs in the related condition compared to the unrelated condition (Table 1), only data from 3 participants remained and there was too little data on which to perform statistical analyses. Fig. 2 provides the average means for children in the different conditions who had more than 33% usable data.

## 4. Discussion

The purpose of this study was to examine the effects of gamification on 2- to 4-year-old children's participation in an experimental study by comparing two versions of a PPNT: a gamified version which included narrative and goal-directed elements of game design, and a more traditional control version. Children in the gamified condition completed more trials and contributed more usable data than children in the control condition. In the game condition, there was no correlation between participants' age and the number of usable trials. The fact that young children in the gamified condition completed as many trials as older children suggests that younger participants had increased attention and motivation; resulting in similar performance to older children with respect to the amount of usable data they contributed. There was a weak positive correlation in the control condition, with younger participants having fewer usable data than older children.

Another goal of the study was to examine whether the quality of the data was influenced by gamification of the PPNT. Although children in the game condition completed more trials, the speed of their responses did not change over the course of the experiment. Moreover, children in the game and control versions did not have different speech responses, as measured in the milliseconds it took them to name the images. This is in line with Hawkins et al. [49]'s findings with adults that gamified tasks had no impact on the quality of data. The current results demonstrate that gamification is a successful strategy for improving children's participation in experimental research, supporting the previous findings of Brewer et al. [15] and Polišenská and Kapalková [48].

The last goal of the study was to determine whether gamification of the PPNT would result in a research methodology that could provide insights into preschoolers' on-line processing of language. It is possible that gamification of the task would result in larger amounts of usable data, but that the task would not reveal anything about the mental processing of language development. The results revealed an effect of priming condition on preschoolers' speech reaction times. These results are in line with established priming research with older children [21,28,29,31–33]. However, the direction of the priming effect was the opposite of what has been previously found in the literature. In the current study, preschoolers had slower SRTs in the related condition (prime 'ball', target 'BOOK') compared to the unrelated condition (primed 'cake', target 'BOOK'). This is called inhibition priming because hearing a similar sounding word slows down or inhibits responses. But in previous studies, children had faster SRTs in the related conditions or what is termed facilitative priming. Although 2- to 4-year-olds in the current study showed a different direction of the priming effect, this is not surprising given that priming tasks are sensitive to slight timing differences in the presentation of the prime and target. We have been continuing this research and have found that when adults are tested on the same experiment, they also show an inhibitory priming effect. Many studies in the literature have found that by manipulating the timing of the stimuli, the direction of the priming effect shifts (e.g., [23,24]). Crucially, this research demonstrates for the first time that the PPNT can be successfully implemented with preschoolers, but only in the game version of the PPNT.

How might the current findings be applied to Human Computer Interaction (HCI) and more specifically Child–Computer Interaction (CCI)? Central to the field of CCI is the development of interactive techniques that are suited to children's capacities [53]. Existing language-based experimental designs that are used with adults (i.e., the basic control version) are potentially overwhelming to young children who are not used to using this type of computer interface [54]. Incorporating basic game-design elements may have made children feel more comfortable with the task and therefore improved their motivation and attention. Children in the current study interacted by providing a verbal response to a prompt (compared to more traditional interactions of touching a screen or a button). Part of the success of our design was that we used a visual cue to indicate to children when they were to give a verbal response. In the gamified condition, Pingu pointed to indicate that children should name the image on the screen and in the control condition, a purple arrow pointed to the target image.

To date, research on CCI has not focused on children's speech as a measurement or as an interactive tool, though some research has looked at the use of CCI in storytelling, which impacts the development of children's oral language skills [35]. The incorporation of verbal interactions into education programs or therapy can provide a way to enhance children's learning. For example, speech-controlled games are increasingly being used as clinical tools to assist speech-language therapists (Apps for Speech-Language Pathology and Practice) [55]. In these tasks, automatic speech recognition systems enable players to control the gameplay through their own speech. The effect of verbal production on learning can be captured in a theory of embodied cognition [56]. Theories of embodied cognition can "explain how children (and adults) learn new movements through imitation ([56]: 34)". However, it is important to note that are limits to the efficacy of verbal interactions on learning. In another line of research in our lab, we have been exploring the role of verbalization on learning. Specifically, we have been investigating whether children are better at learning words during a training phase in two conditions: a speech production condition where children produce the words, and a heard-only condition where children simply hear words and

**Table 1**
Mean SRTs in gamified and control conditions on priming conditions. Standard deviations in parentheses.

| Condition | Age in months | Related condition | Unrelated condition | Difference |
|---|---|---|---|---|
| Gamified, $n = 7$ | 37.61 | 1199 ms (129 ms) | 1082 ms (207 ms) | 117 ms |
| Control, $n = 3$ | 44.63 | 940 ms (309 ms) | 848 ms (194 ms) | 92 ms |

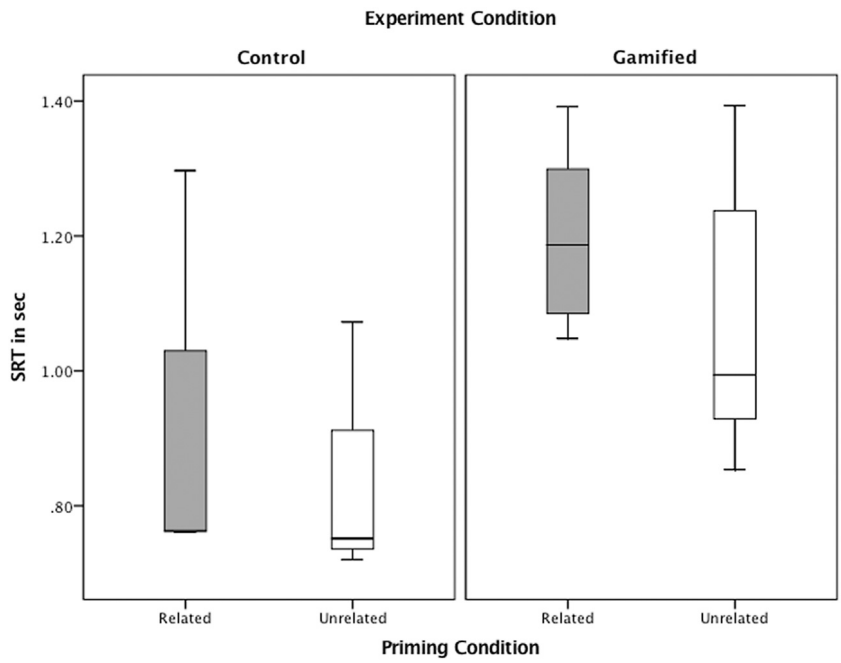


**Fig. 2.** Mean Speech Reaction Times (SRTS) for participants in control and gamified conditions, for responses in Related and Unrelated Priming Conditions. Based on subset of participants with over 33% usable data.

do not provide a verbal response. In some cases, speech production provides a beneficial effect on learning: children recall more words from a training phase that are produced rather than heard-only [57]. However, in a more complex task, the opposite effect is found. In this case, learning is worse in the speech production condition and better in the heard-only condition where children do not provide a verbal response [58]. Thus, an important observation to keep in mind is that the benefits of verbal interactions can be subject to attentional-, linguistic- and task-related demands [59]. Thus, CCI must be flexible to differentiate between situations where verbal interactions can enhance learning and recognize that in other situations, children's learning may be improved if verbal interactions are delayed.

This study illustrates that gamification is a viable strategy for increasing preschoolers' participation on an experimental task and for improving research on language development. Both the length of children's participation and the quality of data were improved in the gamified version. There are many different elements of game design that can be implemented in non-game settings. Deterding et al. [9,60] identifies a hierarchy of game design elements, starting with the most basic elements of game design interface, such as incorporating badges and leadership levels. Future research may focus on which aspects of game-design elements are most useful for the gamification of an experimental task [60], and which elements improve the length and quality of task performance, providing children with a more enjoyable research experience. Although children were not asked about how they felt about the PPNT, the willingness of all children in the gamified version to complete all 40 trials indicates that they enjoyed the experience. With increased interest and improved attention, the expectation is that gamification of experiments with added complexity could also be run with younger children. This in turn would allow for more equivalent comparisons of research across the developmental lifespan. In addition, improving experimental performance and reducing data loss means that fewer participants are needed. In sum, gamification is a useful tool for improving young children's motivation and engagement in experimental research.

## Acknowledgments

This research was supported by a grant from the National Science and Engineering Research Council (Grant No. 2014-05690) of Canada awarded to TSZ, and an MA Fellowship from Social Sciences and Humanities Research Council (Grant No. 766-2013-4211) of Canada awarded to MW.

## References

[1] E.J. Johnson, T.S. Zamuner, Using infant and toddler testing methods in language acquisition research, in: E. Blom, S. Unsworth (Eds.), Experimental Methods in Language Acquisition Research, John Benjamins, Amsterdam, 2010, pp. 73–93.
[2] V. Slaughter, T. Suddendorf, Participant loss due to "fussiness" in infant visual paradigms: A review of the last 20 years, Infant Behav. Develop. 30 (3) (2007) 505–514.
[3] T.S. Zamuner, E.J. Johnson, Methodology in phonological acquisition: Assessing the joint development of speech perception and production, in: B. Botma, N. Kula, K. Nasukawa (Eds.), Continuum HandBook To Phonology, Continuum, Amsterdam, 2011, pp. 16–29.
[4] A. Christophe, S. Peperkamp, C. Pallier, E. Block, J. Mehler, Phonological phrase boundaries constrain lexical access I. Adult data, J. Mem. Lang. 51 (4) (2004) 523–547.
[5] A. Gout, A. Christophe, J.L. Morgan, Phonological phrase boundaries constrain lexical access II. Infant data, J. Mem. Lang. 51 (4) (2004) 548–567.
[6] S. Curtin, K. Byers-Heinlein, J.F. Werker, Bilingual beginnings as a lens for theory development: PRIMIR in focus, J. Phonetics 39 (4) (2011) 492–504.
[7] J.F. Werker, S. Curtin, PRIMIR: A developmental framework of infant speech processing, Lang. Learn. Develop. 1 (2) (2005) 197–234.

[8] R.N. Aslin, J. Fiser, Methodological challenges for understanding cognitive development in infants, Trends Cogn. Sci. 9 (3) (2005) 92–98.
[9] S. Deterding, D. Dixon, R. Khaled, L. Nacke, From game design elements to gamefulness: defining gamification, in: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, ACM, 2011, pp. 9–15.
[10] T.S. Zamuner, C. Moore, F. Desmeules-Trudel, Toddler's sensitivity to within-word coarticulation during spoken word recognition: Developmental differences in lexical competition, J. Exp. Child Psychol. 152 (2016) 136–148.
[11] T.S. Zamuner, Perceptual evidence for young children's developing knowledge of phonotactic probabilities, Language Acquisit. 20 (2013) 241–253.
[12] K. Skoruppa, N. Mani, S. Peperkamp, Toddlers' processing of phonological alternations: Early compensation for assimilation in English and French, Child Develop. 84 (1) (2013) 313–330.
[13] B. Ambridge, .C.F. Rowland, Experimental methods in studying child language acquisition, Wiley Interdiscip. Rev.: Cogn. Sci. 4 (2) (2013) 149–168.
[14] R.B. Miller, C.S. Hollist, Attrition bias, in: N.J. Salkind (Ed.), Encyclopedia of Measurement and Statistics, Volume 1, SAGE Publications, Thousand Oaks, CA, 2007, pp. 57–60.
[15] R. Brewer, L. Anthony, Q. Brown, G. Irwin, J. Nias, B. Tate, (2013) Using gamification to motivate children to complete empirical studies in lab environments, in: Proceedings of the 12th International Conference on Interaction Design and Children, New York, NY, 2013, pp. 388–391.
[16] S. Fisher, Identifying video game addiction in children and adolescents, Addictive Behav. 19 (5) (1994) 545–553.
[17] M.C. Frank, E. Sugarman, A.C. Horowitz, M.L. Lewis, D. Yurovsky, Using tablets to collect data from young children, J. Cogn. Develop. 17 (1) (2016) 1–17.
[18] M. Ota, S.J. Green, Input frequency and lexical variability in phonological development: A survival analysis of word-initial cluster production, J. Child Lang. 40 (2013) 539–566.
[19] T.S. Zamuner, The structure and nature of phonological neighbourhoods in children's early lexicons, J. Child Lang. 36 (2009) 3–21.
[20] B. Munson, A.L. Baylis, M.O. Krause, D. Yim, Representation and access in phonological impairment, in: C. Fougeron, B. Kühnert, M. D'Imperio, N. Vallée (Eds.), Laboratory Phonology, Vol. 10, Mouton de Gruyter, Berlin, 2010, pp. 381–404.
[21] J.D. Anderson, Age of acquisition and repetition priming effects on picture naming of children who do and do not stutter, J. Fluency Disorders 33 (2) (2008) 135–155.
[22] J. Santiago, Implicit priming of picture naming: A theoretical and methodological note on implicit priming task, Psicol. Rev. Metodol. Psicol. Exp. 21 (1) (2000) 39–60.
[23] P.J. Brooks, B. MacWhinney, Phonological priming in children's picture naming, J. Child Language 27 (2) (2000) 335–366.
[24] Q. Zhang, C. Feng, X. Zhu, C. Wang, Transforming semantic interference into facilitation in a picture–word interference task, Appl. Psycholinguist. (2015) 1–25.
[25] A.S. Meyer, H. Schriefers, Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli, J. Exp. Psychol. 17 (6) (1991) 1146–1160.
[26] J. Sadat, C.D. Martin, A. Costa, F. Alario, Reconciling phonological neighborhood effects in speech production through single trial analysis, Cogn. Psychol. 68 (2014) 33–58.
[27] F. Xavier Alario, J. Segui, L. Ferrand, Semantic and associative priming in picture naming, Quart. J. Exp. Psychol. 53 (3) (2000) 741–764.
[28] K.S. Melnick, E. Conture, R.N. Ohde, Phonological priming in picture naming of young children who stutter, J. Speech Lang. Hear. Res. 46 (2003) 1428–1443.
[29] C.T. Byrd, E.G. Conture, R.N. Ohde, Phonological priming in young children who stutter: Holistic versus incremental processing, Am. J. Speech-Lang. Pathol. 16 (1) (2007) 43–53.
[30] L.M. Henderson, P.J. Clarke, M.J. Snowling, Accessing and selecting meaning in autism spectrum disorder, J. Child Psychol. Psychiatry 52 (9) (2011) 964–973.
[31] B. Munson, C.L. Swenson, S.C. Manthei, Lexical and phonological organization in children: evidence from repetition tasks, J. Speech Lang. Hear. Res. 48 (1) (2005) 108–123.
[32] M. Pellowski, E. Conture, Lexical priming in picture naming of young children who do and do not stutter, J. Speech Lang. Hear. Res. 48 (2005) 278–294.
[33] C. Savage, P. Howell, Lexical priming of content and function words with children who do, and do not, stutter, J. Commun. Disorders 41 (6) (2008) 459–484.
[34] D.K. Dickinson, A. McCabe, L. Anastasopoulos, E.S. Peisner-Feinberg, M.D. Poe, The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children, J. Educ. Psychol. 95 (3) (2003) 465.
[35] C. Sylla, C. Coutinho, P. Branco, M. Müller, Investigating the use of digital manipulatives for storytelling in pre-school, Int. J. Child-Comput. Interact. 6 (2015) 39–48.
[36] S. Deterding, Gamification: Designing for motivation, Interactions 19 (4) (2012) 14–17.
[37] T.W. Malone, Toward a theory of intrinsically motivating instruction, Cogn. Sci. 4 (1981) 333–369.
[38] R. Rosas, M. Nussbaum, P. Cumsille, V. Marianov, M. Correa, P. Flores, V. Grau, F. Lagos, V. Lopez, P. Rodriguez, M. Salinas, Beyond Nintendo: Design and assessment of educational video games for first and second grade students, Comput. Educ. 40 (2003) 71–94.
[39] M.J. Dondlinger, Educational video game design: A review of the literature, J. Appl. Educ. Technol. 4 (1) (2007) 21–31.
[40] J.P. Gee, Learning by design: Good video games as learning machines, E-Learning 2 (1) (2005) 5–16.
[41] A. Amory, K. Naicker, J. Vincent, C. Adams, The use of computer games as an educational tool: Identification of appropriate game types and game elements, Br. J. Educ. Technol. 30 (4) (1999) 311–321.
[42] BioWare, Dragon Age, Electronic Arts, Redwood City, CA, 2009.
[43] Rovio, Angry Birds, Rovio Entertainment, Espoo, Finland, 2009.
[44] A. Waraich, Using narrative as a motivating device to teach binary arithmetic and logic gates, ACM SIGCSE Bull. 36 (3) (2004) 97–101.
[45] A. Achim, A. Marquis, An automated version of the BAT Syntactic Comprehension task for assessing auditory L2 proficiency in healthy adults, Clin. Linguist. Phonetics 25 (6–7) (2011) 655–669.
[46] A. Marquis, (2013) Using Android applications in research: Introducing AndroidBAT and Jeu de verbes. Paper presented at the 2013 PsychoShorts Conference, Ottawa, Canada.
[47] A. Bello, P. Giannantoni, P. Pettenati, S. Stefanini, M.C. Caselli, Assessing lexicon: Validation and developmental data of the Picture Naming Game (PiNG), a new picture naming task for toddlers, Int. J. Lang. Commun. Disorders 47 (5) (2012) 589–602.
[48] K. Polišenská, S. Kapalková, Improving child compliance on a computer administered nonword repetition task, J. Speech Lang. Hear. Res. 57 (3) (2014) 1060–1068.
[49] G.E. Hawkins, B. Rae, K.V. Nesbitt, S.D. Brown, Gamelike features might not improve data, Behav. Res. Methods 45 (2) (2013) 301–318.
[50] P. Boersma, D. Weenink, (2010) {P}raat: doing phonetics by computer, version 5.3.79. [Computer program].
[51] SR Research, Experiment Builder [Version 1.10.1025], SR Research Ltd, Ottawa, 2013.
[52] M.D. Tyler, L. Tyler, D.K. Burnham, The delayed trigger voice key: An improved analogue voice key for psycholinguistic research, Behav. Res. Methods 37 (2005) 139–147.
[53] J.C. Read, P. Markopoulos, Child–computer interaction, Int. J. Child-Comput. Interact. 1 (2013) 2–6.
[54] L. Plowman, J. McPake, Seven myths about young children and technology, Childhood Educ. 89 (1) (2013) 27–33.
[55] Apps for Speech-Language Pathology and Practice. From http://www.asha.org/SLP/schools/Applications-for-Speech-Language-Pathology-Practice/.
[56] A. Antle, Research opportunities: Embodied child–computer interaction, Int. J. Child-Comput. Interact. 1 (2013) 30–36.
[57] M. Icht, Y. Mama, The production effect in memory: A prominent mnemonic in children, J. Child Lang. 42 (2015) 1102–1124.
[58] T.S. Zamuner, S. Strahm, E. Morin-Lessard, M.P. Page, Reverse production effect: Children recognize novel words better when they are heard rather than produced during training. Developmental Science (in press).
[59] T.S. Zamuner, H.H. Yeung, M. Dumos, The many facets of speech production and its complex effects on phonological processing, Br. J. Psychol. 108 (2017) 37–39.
[60] S. Deterding, M. Sicart, L. Nacke, K. O'Hara, D. Dixon, Gamification. using game-design elements in non-gaming contexts, in: CHI'11 Extended Abstracts on Human Factors in Computing Systems, ACM, 2011, pp. 2425–2428.