Many interventions have been developed to address reading difficulties in clinical (e.g., Specific Learning Disabilities) and the general population of struggling readers and these have shown promise in research. However, these interventions have had a very limited impact on psychological practice as most psychologists do not implement psychoeducational interventions, but just provide recommendations for teachers, who become solely responsible for addressing the struggling reader’s needs. A plausible explanation for these gaps between research and practice is limitations in access to these interventions by psychologists and by the populations that need these services. When the literature on the efficacy of these interventions is examined closely many limitations are easily identified. For example, most interventions require specialized training (e.g., doctoral, postdoctoral studies), and are expensive and difficult to implement. Furthermore, their efficacy is usually demonstrated under very rigorous implementation conditions such as long sessions and very frequent weekly sessions (e.g., four –five times) that span a few months.

When the conditions described by these studies are compared to real-world scenarios, it seems evident why these interventions have failed to impact psychological practice. For example, in the context of Puerto Rico, health insurance usually does not cover the expenses of psychological services, particularly in the educational fields. It should not be surprising then, that most struggling readers will not be able to receive psychological services four to five times a week and in most cases the frequency and consistency of these interventions are important in order to achieve improvements.

These limitations of psychological and educational services in Puerto Rico highlights the importance of developing more flexible and viable alternatives to address the difficulties of struggling readers. The integration of technology to psychological and educational practice has been a venue by which to address these limitations, as evidenced by the increasing number of studies addressing the efficacy of technology-based learning interventions, particularly for basic academic skills, such as reading and mathematics. It is important, however, to investigate if technology-based interventions are effective and if these address the limitations that traditional interventions have encountered when considered in real-world scenarios. The present systematic review of the literature analyzes studies that test the effects of technology-based interventions on reading skills to address the following specific aims: (a) identify which cognitive processes mediate the impact these interventions have on reading skills and which seem to be more effective, (b) identify common research methods and instruments employed in these studies (e.g., design, sampling procedures), (c) identify common analytical methods employed in these studies (e.g., inferential statistics, effect sizes).

# Methods

A detailed description of the criteria used to determine the studies eligibility and the protocol implemented to select these is presented below in order to guarantee the replicability of the findings. These were established *a priori,* as is customary when conducting systematic literature reviews.

## Selection criteria

In order to identify relevant and recent literature the following criteria must have been met by the articles in order to be selected: (a) published between 2009 – 2019 (i.e., last ten years), (b) published in a peer-reviewed journal, (c) consist of an empirical and quantitative study (e.g., journal article, thesis, or dissertation), (d) the intervention must be administered through a computer, tablet or other technological device, (e) the intervention must have addressed one or more basic reading skills (e.g., decoding, word recognition, word reading, phonological awareness, phonemic awareness, phonics, spelling), and (f) the participants assisted elementary school.

## Search protocol

The consulted database providers and databases are presented in table 1. These were selected as they are pertinent to the field of psychology and education. The keywords used during the search are presented in table 2 and were selected in order to identify studies relevant to the population of interest, that used technology-based interventions, and that focused on basic reading skills. Whenever available, the delimiters for using *related keywords*, *equivalent subjects*, and *scholarly/peer reviewed* were selected.

Table 1.

Database providers and databases.

| Database provider | Databases |
| --- | --- |
| Academic Search Ultimate | Academic Search Ultimate, E-Journals, ERIC, Fuente Académica Plus, MEDLINE with Full Text, Psychology and Behavioral Sciences Collection, PsycINFO, Teacher Reference Center |
|  |  |
| ScienceDirect | Across all databases (general search) |

Table 2.

Keywords used during the search process.

| Keyword combinations | | | |
| --- | --- | --- | --- |
| Elementary | Computer\* | Reading | Intervention |
| Elementary | Computer\* | Reading | Remediation program |
| Elementary | Computer\* | Reading | Phonics |
| Elementary | Computer\* | Reading | phonological awareness |
| Elementary | Computer\* | Reading | decoding |
| Elementary | Computer\* | Reading | word Recognition |
| Elementary | Computer\* | Reading | phonemic Awareness |

Note: \*refers to the wild card character used to match any keyword starting with the string  
computer (e.g., computerized, computer).

## Study Selection

In order to identify the studies the following procedure was followed: (a) searches were conducted on the relevant database; (b) the study’s title and abstract was reviewed to determine its eligibility; (c) eligible articles were downloaded; (d) duplicates were removed; (e) the study’s full text, with a particular emphasis on its title, abstract and methods section was reviewed in order to determine its eligibility. During the last step of the process, four judges reviewed each article to determine its eligibility. Whenever there was disagreement between the judges, it was discussed and resolved. The articles presented were deemed eligible by all authors. Table 3 presents the screening process.

Table 3

Screening process for articles.

| Phase | Results | Reasons for discarding articles |
| --- | --- | --- |
| Initial search | 451 | - |
| Initial screening | 29 | Older children, reading skills were not measured, non-empirical article, severe neurodevelopmental disorders (e.g., autism spectrum disorder), non-technological-based intervention, not focused on basic reading skills, qualitative studies, focused on academic skills other than reading |
| Duplicate removal | 25 |  |
| Final article selection by judges | 21 | Did not address basic reading skills, the same study was published with different titles and variations of their analytical approacha |

Note. aOnly one study was included in this review.

## Coding

Some information about the studies had to be coded in order to analyze the information. Table 4 shows the categories and their definitions. Only categories that were coded or that include non-obvious subcategories are presented for brevity and simplification.

Table 4.

Definitions for categories

| Dimension | Category | Sub-category | Definition |
| --- | --- | --- | --- |
| Methods | Design | Single Subject Design (SSD) | Any variant of SSD were coded as SSD |
| Pretest-posttest design with comparison group(s) | Design that included pretest, posttest, and one comparison group |
| Pretest-posttest design with multiple comparison groups | All design that included pretest, posttest, and multiple comparison groups (e.g., multiple baseline profiles) |
| Pretest-posttest design with multiple experimental groups | All design that included pretest, posttest, and multiple experimental groups (e.g., various technology-based interventions were tested) |
| Methods and findings | Reading skills | Phonological skills | Skills that required phonological processing (e.g., phonemic awareness, phonological awareness) |
| Word reading | Skills that required reading or recognizing words |
| Phonics | Skills that required the integration or knowledge about grapheme and phoneme |
| Reading: Otherwise not specified | No specific skill was identified |

## Cognitive component

Cognition comprises the underlying mental processes by which we learn, and research has established an intricate and direct and indirect relationship between cognitive processes and reading skills. The original plan for this review was to identify the cognitive processes that reading interventions addressed or considered because from a neuropsychological perspective, this provides the best rationale for designing academic interventions (Feifer, Kaufman, & Kaufman, 2011). However, only one study was identified which addresses cognitive processes in the process of improving reading skills through intervention. For this reason, the results of this study will be complemented by a discussion of how cognitive processes are related to reading skills with the aim of providing information that will aid in the development of reading interventions that take into consideration the cognitive mechanisms that underlie these.

The study by Messer et al (2008) tested the effects of the Trainertext intervention on working memory and phonological short-term memory and basic reading skills. The findings show a medium effect (d = .65 - .69) on these cognitive processes and a small to large effect (d = .27 - .97) on reading fluency and spelling. The findings suggest that the intervention, efficacious at improving working memory, phonological short-term memory, reading fluency, and spelling skills.

Research on cognition and reading has identified several cognitive processes that are vital to the development of reading skills. This research and how these particular cognitive processes support reading is discussed. Basic reading skills (e.g., phonological skills, spelling) are associated with various cognitive processes. For example, phonological mechanisms of working memory (WM) are associated with dyslexia (Beneventi et al., 2010; Cruz-Rodrigues et al., 2014; Pinto & Peixoto, 2011). These mechanisms support the reading process by providing the ability to manipulate the phonological elements of language (e.g., producing the sounds associated with particular letters and integrating these sounds into words). This relationship has proven particularly useful at understanding reading difficulties because it provides a parsimonious explanation of the difficulties presented by these children. Visual mechanisms of working memory have also been associated with reading skills (Pinto y Peixoto, 2011) because the inefficient processing of visual representations of sounds (i.e., graphemes) could disrupt the integration of visual and auditory components of written text, resulting in poor reading skills.

Higher order memory processes also support reading skills. Cruz-Rodrigues et al. (2014) have found semantic memory deficits in children who experience reading difficulties. Semantic memory is involved in the coding of general knowledge (i.e., information that lacks context; cite Groome, 20114) and could be associated with difficulties in retrieving information about a specific sounds (i.e., phoneme) that corresponds to a letter (i.e., grapheme) or word and the retrieval of the meaning of words that is often experienced by struggling readers. Furthermore, deficits in successive processing of information has also been linked to reading difficulties (Keat & Hj. Ismail, 2011). Successive processing of information involves the sequential processing of information (i.e., ordered) and could help explain the reason why many struggling readers have severe difficulties in pronouncing words correctly (Das, Naglieri, & Kirby, 1994).

Attention is another cognitive process that has been associated with reading skills. Studies have shown a relationship between phonological awareness and divided attention and alertness (Lewandowska et al., 2014). These difficulties could be related to a reduced ability to focus on the phonological components of letters and words, which would lead to errors in the retrieval of the sounds and their integration. These phonological processing problems have also been related to deficits in processing speed and verbal fluency (Moura et al., 2015). In these cases, the loss of speed in the decision-making process of selecting the appropriate sound for a given grapheme may result in difficulties integrating sounds into words and could lead to a slow and fragmented reading process. This would in turn result in limited verbal fluency because the speed with which words are retrieved based on their meaning or phonological characteristics is limited.

Another area of interest are complex reading skills. Similar to basic reading skills, complex reading skills have been associated with various cognitive processes. For example, simultaneous processing of information (i.e., integration of information into a whole) deficits have been found in struggling readers (Keat & Hj. Ismail, 2011) and these deficits could explain their difficulties in deriving meaning from words and sentences (Das et al., 1994). Attention has also been associated with reading comprehension, although its relationship to basic reading skills is stronger (Keat & Hj. Ismail, 2011).

With regards to reading comprehension, deficits in attention are believed to be associated with loss of information and this is detrimental to comprehending the narrative that is being read. Complex reading skills have also been associated with executive functions (FE) deficits. Children with Dyslexia have been found to present deficits in efficiently regulating their focus of attention (Krause, 2015; Lewandowska et al., 2014; Moura et al., 2015; Pinto & Peixoto, 2011) and response inhibition (Pinto & Peixoto, 2011). These deficits are believed to limit children’s ability to change their attention between the different parts of the text (i.e., ideas) in order to achieve an integrated picture of the narrative (i.e., derive a central idea). Finally, deficits in cognitive flexibility (Lewandowska et al., 2014) and planning (Keat & Hj. Ismail, 2011) have been found in children with dyslexia. Deficits in these areas could explain the reason why struggling readers do not benefit as expected from strategies to improve reading comprehension or make the necessary adjustments to the information that has just been read in order to integrate the information that will be read.

It is important to note that although the same cognitive processes are often associated with different reading skills (e.g., attention is associated with basic and complex reading skills), there are differences in how these support reading skills. This fact should not be surprising because cognition, as well as the brain, works as an integrated system by which we perceive the world and process information. For most behaviors, we use an intricate array of cognitive processes and brain circuits, but certain processes and circuits are more involved than others for any given behavior. This same concept applies to reading skills. During the reading process, an intricate array of cognitive processes and brain circuits are involved. Some may change during different phases of the reading process (e.g., cognitive planning) and other may always be involved to the same extent (e.g., attention). This evidences the complexity of designing interventions for improving reading skills derived from the cognitive processes that are associated with them as these intricate relationships must be considered.

## Methodological component

The studies reviewed primarily used pretest-posttest design with multiple experimental groups (33.33%), pretest-posttest design with comparison group (23.81%), and randomized control trials (RCT) with pretest-posttest (19.05%). The term *multiple experimental groups* was used to identify studies in which different groups were exposed to different technology-based interventions and the groups improvements were compared. The study design had to be inferred for 19.05% of the articles discussed because these did not state an explicit study design. In these cases, the study design was inferred from more general descriptions in the methods section (e.g., random assignment, administration of pretest and posttest). Most studies used random assignment (66.67%) to allocate participants to each group in the study, yet not many were considered by their authors to meet the standards to be named RCTs. Studies were balanced with regards to balancing the different study’s groups based on important characteristics (balanced = 46.67%, not balanced = 53.33%), but only a few studies counterbalanced the order of test administration (16.67%), which is extremely important to protect the validity of the measurements, or used a probabilistic sampling (9.52%). The median number of participants across studies was 31, but it varied greatly across studies (minimum = 2, maximum = 744). Most studies worked with four- to eight- year-old children (66.67%) from Kindergarten to fourth grade (66.67%), whose main language was English.

The described studies present various strengths and limitations. These implemented strong designs with pretest-posttest and comparison groups and used random assignment to balanced participant characteristics across the groups. However, must studies do not counterbalanced testing procedures, which may have resulted in patterns of participant performance related to the order of administration and not to differences between groups. Additionally, studies did not use probabilistic sampling schemes, and this makes it difficult to determine the generalizability of the findings to population. The sample size varied greatly across studies but in general, these consisted of small to medium studies. The sample characteristics are important as these helps to understand the population to which these results may generalize. In general, the fact that most interventions were tested on younger children in the elementary grades (i.e., Kindergarten to fourth grade) is important because it means that the interventions are suitable to be used as early intervention programs. Another important characteristic is the native language in which the intervention was administered. Most studies focused on reading skills of English and French languages which are considered as having an opaque orthography because the same sound (i.e., phoneme) can be associated with different visual representations (i.e., graphemes; cita).

Intervention implementation details are also described. A great number of interventions were tested in the different studies, but GraphoGame, in its different variants, is the most used technology-based intervention (33%). Many studies did not report the number of sessions provided to the participants (38.10%), but there was great variability among those that provided the information (i.e., from less than 20 to more than 40 sessions). In some cases, the total amount of time dedicated at working with the interventions was provided. Interventions were administered most commonly four times a week (50%) individually or in groups in 10- to 30-minute sessions (76.19%). There was a general tendency across studies to not report the size of the groups (45.45%), but small groups were preferred (2 – 3 participants; 36.36%). Note that in this review, interventions were considered *group-administered* if more than one participant received the intervention in the same space and time as another participant, even if they worked completely independently. Finally, most interventions were administered in controlled settings (e.g., schools) under supervision (90.48%) but many authors reported only providing support with technical difficulties.

Implementation details provide information about under which circumstances were the interventions tested. This information is very valuable in determining the viability of the interventions in real-word scenarios. It seems there are many alternatives in technology-based interventions, as evidenced by the great number of interventions used. In most cases, the number of sessions were not provided, which imposes great limitations to determining how many sessions are necessary to improve academic skills. The fact that the interventions were implemented 4 times a week is less than ideal, as it imposes a challenge to expose children that many times a week to a supervised intervention, even if it is computerized and mobile. However, short sessions were more common and being able to administer the interventions in small groups contributes to their viability because many students can work in a single session and the investigators mostly provided only technical support and children worked individually.

## Statistical component

In this section, the statistical components of these studies are explored. A detailed description of the analytical methods used in studies is extremely important as these helps to determine the validity of the findings in any study. Most of the studies reported descriptive analysis results of the main outcomes (90.48%). Descriptive statistics are important and describe the variables of interest and allow to determine if any strange patters were observed in the data, which may indicate problems with sampling (e.g., mean scores that are too low may indicate that the control group may not be a *typically developing group)*. The fact that descriptive statistics were provided in most studies. The mean or median (90%), standard deviation (75%) were frequently reported, but only a few studies did report confidence interval for the means (11.76%). Mean confidence intervals are extremely important for estimating the true mean of the population, but are usually not emphasized in favor of the predominant p-value and point-estimate approach. This omission of confidence intervals comes at the expense of losing confidence in the certainty of the estimated true value of the population mean, as a point-estimate (e.g., estimated population mean from the mean of the sample) is usually not very reliable across studies. Other reported descriptive statistics are total score and accuracy, these were reported in 33.33% and 50% of the reviewed studies, respectively.

With regards to inferential statistics, ANCOVA (38.89%) and ANOVA (33.33%) were the two most popular choices among studies. Baseline reading skills was the most commonly controlled variable across studies that controlled any variables (38.89%), but most studies did not control any variables (44.44%). Only one study controlled for variables other than cognitive ability or academic skills (i.e., age, income, mother’s age, mother’s education). Another important detail to note is that most studies did provide specific p-values for the analysis conducted (77.78%), particularly when statistical significance was not reached. The actual *p-value* is important, contrary to popular believe, as this value should not be interpreted as a binary outcome (i.e., statistically significant or nonsignificant), but as the probability of findings results as extreme as the ones found, if there was no difference. This means that p-values can suggest whether a *nonsignificant* result is still worth pursuing in further research or if it is unlikely there is a true relationship among the variables studied.

Surprisingly, most studies did provide effect sizes (88.89%) and the most commonly reported was Cohen’s d (56.25%). Effect sizes are an indicator of the magnitude of the relationship (i.e., in this case, the magnitude of the intervention effect) and should be reported independently of statistical significance (cite Durlak). Consistent with the reporting of confidence intervals for the mean, only a few studies (14.29%) reported the confidence interval for the mean difference (i.e., differences between the means).

The studies reported follow some of the best practices in the report of results. These use sound analytical methods (e.g., ANOVA, ANCOVA), provide descriptive statistics, and provide the specific p-values and effect sizes. Additionally, the studies controlled for variables that are known to influence improvements after the implementation of an intervention (e.g., baseline reading skills). However, the studies could have presented Hedge’s G as the effect size index, which as has been described as an unbiased estimator of the true effect, and corrections appropriate for small sample sizes could have been used to avoid overestimating the intervention’s effect. Confidence intervals were another very important statistic that is missing in most of the reviewed studies. As stated previously, mean confidence intervals allow better estimation of the population means and mean difference confidence intervals allow to estimate the true difference between the groups as an interval that is more reliable across studies.

## Study findings component

The most commonly addressed reading skills were phonological skills (e.g., phonological awareness, phonemic awareness; 26.79%) and word reading (25%). Across the studies, word reading and phonological skills were the ones that showed the greatest improvement (29.73%) and (16.22%), respectively. The effect sizes evidenced medium (68.75%) and large (81.25%) intervention effects across all reading skills that showed improvements across all studies. The classifications of the effect sizes were interpreted following the general guidelines provided by Cohen (1988). Many studies reported interventions having an effect on all targeted skills (33.33%), but most reported not being able to improve at least one reading skill. Phonological skills were also the reading skill that most interventions failed to show improvements on (16.67%). This apparent contradiction is due to how the original reported reading skills were coded. Most studies targeted one or more components of phonological skills and some of these components showed improvements while others did not.

The effect of these interventions on basic reading skills such as phonological skills and word reading is apparent. It is also important to note that these effects were medium and large in most studies, which is particularly encouraging. The particular skills that participants showed improvements on are consistent with their development level, given they were mostly young children in primary grades. The fact that most interventions failed to have an effect on some of the targeted skills demonstrates inconsistencies in the intervention’s efficacy. It seems these are better at improving some basic reading skills that others, particularly within the domain of phonological skills.

## Conclusion

Technology-based reading interventions are being developed as a complementary tool to address these important skills. The findings of this systematic review suggest there are many different options when it comes to technology-based interventions and these have great potential for improving reading skills. Studies that test these implement strong designs, conduct strong statistical analyses and comply with good statistics reporting standards. Furthermore, these interventions have been tested with short sessions and in groups which suggest these may be viable in real-world scenarios. In the context of Puerto Rico viability is very important as the limited access to resources and the high prevalence of this population (Disdier‐Flores & Jara Castro, 2017) requires cost-effectively tools to address this population’s needs.

The findings also suggest that many studies do not use sampling schemes that allow generalizations to the population. Confidence intervals are also missing, and this is a very important statistics because it would allow to better estimate the improvements caused by the interventions. In this line, the use of Hedges’ G as an index of the size of the intervention’s effect would result in better estimates, as has been previously stated. Future studies should make efforts to include these statistics to strengthen confidence in their results. Additionally, the impact of technology-based interventions on more transparent orthographies (e.g., Spanish and Italian) should be explored. It would also be interesting to compare the specific teaching methodologies the interventions implement and determine which are more effective.

Another aspect that is particularly important but missing from the reviewed studies is the use of knowledge of the relationship between cognition and reading to derive interventions that address the development of reading skills through their cognitive correlates. How cognition and brain circuits are used during the reading process should play a central role in intervention design, planning, and monitoring. One reason why consideration about cognition may be absent from the reviewed interventions is that a deep understanding of how cognition supports reading is required in order to design these interventions because the mechanisms by which cognition support reading are intricate. Nonetheless, future studies should address the cognitive mechanisms by which these interventions impact reading skills.

Despite the limitations discussed, most of the studies reviewed had strong methodological and statistical characteristics and the findings of this review suggest that technology-based interventions have high potential for improving reading skills. More research should be conducted to continuously refine these interventions in order to improve their efficacy and cost-effectiveness. Technology is flexible, powerful, and highly engaging for children which makes it an ideal venue to explore.

# References