

Case Study 1: An Evidence-Based Practice Review Report

How effective are one-to-one therapy dog interventions on reading ability in primary school children?

1. Summary

Primary-aged children's reading ability is a key concern for government and current policy (Department for Education, 2015) with a large number of these children not meeting the expected levels of reading (Department for Education, 2019). Therapy dogs are currently growing in popularity in many countries, including the UK, with many schools already having a therapy dog. There are a number of key differences in the relational aspects of interventions where the primary recipient for the child is an animal rather than a human. Therapy dog interventions in schools can be different in nature, in this case the interventions involve a child reading to a dog without judgement or correction from another person, and in this review this was on a one-to-one basis. This systematic literature review examined five studies which met the inclusion criteria and these were critically appraised for their methodological rigour and relevance. Effect sizes were examined to answer the review question.

There were mixed findings with a number of large, medium and small effect sizes, as well as some studies with no effects observed. There were a number of methodological flaws across the studies, in addition to great variability in the study designs and few participants overall per study. To summarise, caution should be given to concluding that there was an overall positive impact of therapy dog interventions on reading ability in primary schools, and future research is needed to support or refute this notion.

2. Introduction

2.1 What is a therapy dog?

A therapy dog is a trained dog used alongside a trained adult to support a person psychologically, socially, emotionally or physiologically (Friesen, 2010). Therapy dogs have been used with both adults and children in a range of settings such as hospitals and schools (Geist, 2011). Animal-assisted therapy (AAT) as it is known, has been used for many decades, with Levinson being one of the first to coincidentally notice the positive impact of his dog on a number of children's behaviour (Reichert, 1998). "The distinguishing features of Animal-Assisted Therapy (AAT) are characterized by the supplemental inclusion of a trained therapy dog in reaching an intervention goal in therapeutic environments, and as a supplement to an educational objective in school contexts" (Friesen, 2010, pg. 261).

Evidence supports the idea that children find animals comforting (Parshall, 2003), and Geist (2011) suggested that therapy dogs provide an emotional aid to humans during a range of interventions. Therapy dogs have been reported to have a positive effect on physiological stress and anxiety across a range of situations (Friedmann, Thomas & Eddy, 2000). For example, reading to dogs has been shown to reduce children's stress as measured by heart rate and blood pressure (Friedmann, Katcher, Thomas, Lynch & Messent, 1983). Whilst AAT have been reported with a wide range of animals, research suggests that therapy dogs are convenient and cost effective (Friesen, 2010).

Reading to a therapy dog has been documented mostly in schools and libraries where children read to these dogs either on a one-to-one or group basis (Hughes, 2002; Le Roux, Swartz & Swart, 2014). With a growing interest in the intervention, programmes can vary enormously in terms of the structure of the sessions, the session duration

and the number of weeks for the intervention. For example, some sessions allow the child or children to pet the dog, whilst others do not allow any touching of the dogs. Similarly, some interventions do not occur on a regular basis or for a set amount of time, for example this is more likely to be the case in a library group where different people attend on different occasions and it may be less controlled. Another area of difference is that interventions can be used with whole classes, small groups or individuals (Connell, Tepper, Landry & Bennett, 2019; Kirnan, Siminerio & Wong, 2016).

2.2 Psychological Theory

To understand the effectiveness of therapy dogs, principles of social and emotional psychology are fundamental. Many authors describe the relationship between the dog and child during the reading process as unique when compared to a child's relationship with another human (peer or adult) when reading (Friesen, 2010). When reading to a person there are social expectations and judgements made by the child and the adult, as well as the perceived expectations the child thinks adult has (Wohlfarth, Mutschler, Beetz & Schleider, 2014). It is believed that these automatic thoughts that a child processes about the reading interaction that is forthcoming, have an impact on how the child performs during reading, indicated through stress and other physiological measures (Schretzmayer, Kotrschal & Beetz, 2017). Furthermore, whilst a child reads, they are in continual fear of the judgements made by adults and their peers (Wohlfarth et al., 2014). For some children who do not find reading difficult this thought process may be very different and not problematic. However, for children who have additional needs, find reading difficult and struggle with their confidence and self-esteem this reading interaction is not a positive experience. To add to this, Wohlfarth et al. (2014) describe self-worth theory where they state that a child is likely to stop

trying if they expect negative feedback from another person. “Most frequently such negative feedback occurs after an experience of failure. Failure threatens self-estimates of ability and creates uncertainty about an individual’s capability to perform well thereafter. If the subsequent performance turns out to be poor, then doubts concerning one’s own ability are confirmed” (Wohlfarth et al., 2014, pg. 62). A key component of interventions with therapy dogs is that there is no feedback provided to the child whilst they are reading, including any corrections on errors made. This makes the reading experience less critical, with fewer social confinements and overall a more positive experience for the child.

This evidence suggested that providing children with the opportunity to participate in interventions with therapy dogs may give them the opportunity to practise in an environment where they have less anxiety and fear of judgement and negative feedback from humans. As a result, therapy dog interventions may facilitate better reading abilities in children.

2.3 Rationale

Reading is considered a fundamental part of development for children (Department of Children, Schools & Families, 2003). Specifically, reading is a key communicative tool that is central to social, emotional and cognitive abilities (Rose Report, 2006). A report by Save the Children (2015) highlights that in children aged eleven on free school meals, only four out of ten are competent at reading. In the UK almost 30% of children are not achieving the expected levels of reading at Key Stage 2 (Department for Education, 2019). School success is known to impact upon later life outcomes (Johnson & Kosykh, 2008), specifically children who fall behind at school are more likely to experience negative life outcomes, for example relating to health, social welfare and income, as well as engage in crime (Allen, 2011).

As stated in a number of government documents, such as the SEND Code of Practice (DfE & DoH, 2015) and the Every Child Matters agenda (Department of Children, Schools & Families, 2003) ensuring the best outcomes for children and young people is imperative. Educational Psychologists have the opportunity with direct working with children, schools and their families to promote better outcomes for reading through intervention programmes. Additionally, Educational Psychologists in the UK are often involved with suggesting evidence-based reading interventions at the request of schools.

There has been popularity and an increase in the number of dogs that are being used in schools in the UK, as well as across other countries (Jalongo, Astorino & Bombo, 2004). For example, in the UK there are programmes supported by the Bark and Read Foundation (Bark & Read, n.d) where volunteers take dogs into school. Furthermore, the former UK Education Secretary Damian Hinds suggested that he believed dogs in schools were having a positive impact and encouraged more schools having a dog for intervention purposes. With this in mind, a better understanding of what the evidence is to support the use of therapy dog interventions on reading is needed.

2.4 Review Question

How effective are one-to-one therapy dog interventions on reading ability in primary school children?

3. Critical Review of the Evidence

3.1 Literature Search

A systematic literature search was conducted in December 2019 for electronic journals using three databases, PsycInfo, ERIC and Web of Science. PsycInfo was selected for its focus on psychological research, ERIC for its educational focus and Web of Science for a broader range of literature. In addition an ancestral search was performed but no further results were found.

The following search terms were used:

Intervention Type: (dog therapy OR therapy dog OR puppy therapy OR therapy puppy OR therapy canine OR canine therapy OR school dog OR animal-assisted OR animal-assisted reading OR canine-assisted OR canine-assisted reading OR classroom dog OR classroom puppy OR classroom canine)

AND

Population: (primary aged OR primary OR primary-aged OR key stage 1 OR key stage 2 OR key stage one OR key stage two OR elementary OR four year old OR 4 year old OR five year old OR 5 year old OR six year old OR 6 year old OR seven year old OR 7 year old OR eight year old OR 8 year old OR nine year old OR 9 year old OR ten year old OR 10 year old or eleven year old OR 11 year old OR year 1 OR year one OR year 2 OR year two OR year 3 or year three OR year 4 OR year four OR year 5 OR year five OR year 6 OR year six OR kindergarten OR first grade OR second grade OR third grade OR fourth grade OR fifth grade OR sixth grade)

AND

Outcome Measure: (reading OR literacy)

3.2. Inclusion and Exclusion Criteria

The search returned 60 results across the three databases. Screening for duplicates removed nine studies, followed by title and abstract screening against the inclusion and exclusion criteria leaving 11 studies. These 11 studies were screened at full-text level allowing further details to be considered against the inclusion and exclusion criteria and this process resulted in five studies remaining (see Figure 1). A detailed description of the inclusion and exclusion criteria for this can be seen in Table 1. For a full list of the studies included in this review see Table 2. A list of the studies excluded at full text screening can be seen in Appendix A. For a table describing the key features of each study see the Mapping the Field table in Appendix B.

Table 1

List of Inclusion and Exclusion Descriptors by Criteria

Criteria	Inclusion	Exclusion	Rationale
1. Type of publication	Published in a peer-reviewed journal	Not published in peer-reviewed journal	To ensure methodological quality and rigour of the research
2. Language	Published or translated into English	Not published or translated into English	To ensure researcher understanding
3. Intervention	a) A therapy dog reading intervention b) Delivered on a one-to-one basis c) Intervention was more than one session	a) Interventions without a therapy dog reading component b) A therapy dog reading intervention in groups c) No intervention, therapy dog reading to assess casual impact (one session)	The review focuses on the impact of a therapy dog reading intervention on a one-to-one basis
4. Participants	Attending a primary school	Not attending a primary school	The review focuses on primary children due to children learning to

Criteria	Inclusion	Exclusion	Rationale
			read and their reading developing fluency at this age
5. Design	<ul style="list-style-type: none"> a) Quasi-experimental b) Randomised Control Trial c) Experimental design (pre and post test data) d) Primary data e) Quantitative data 	<ul style="list-style-type: none"> a) Non-experimental design b) Secondary data c) Qualitative data 	The review aims to assess casual impact
6. Date of publication	2006 – 2019	Pre 2006	Considers literature since the implementation of the phonics review (Rose Report, 2006) as a strategy for teaching reading

Figure 1

Flow Diagram of Selection Process

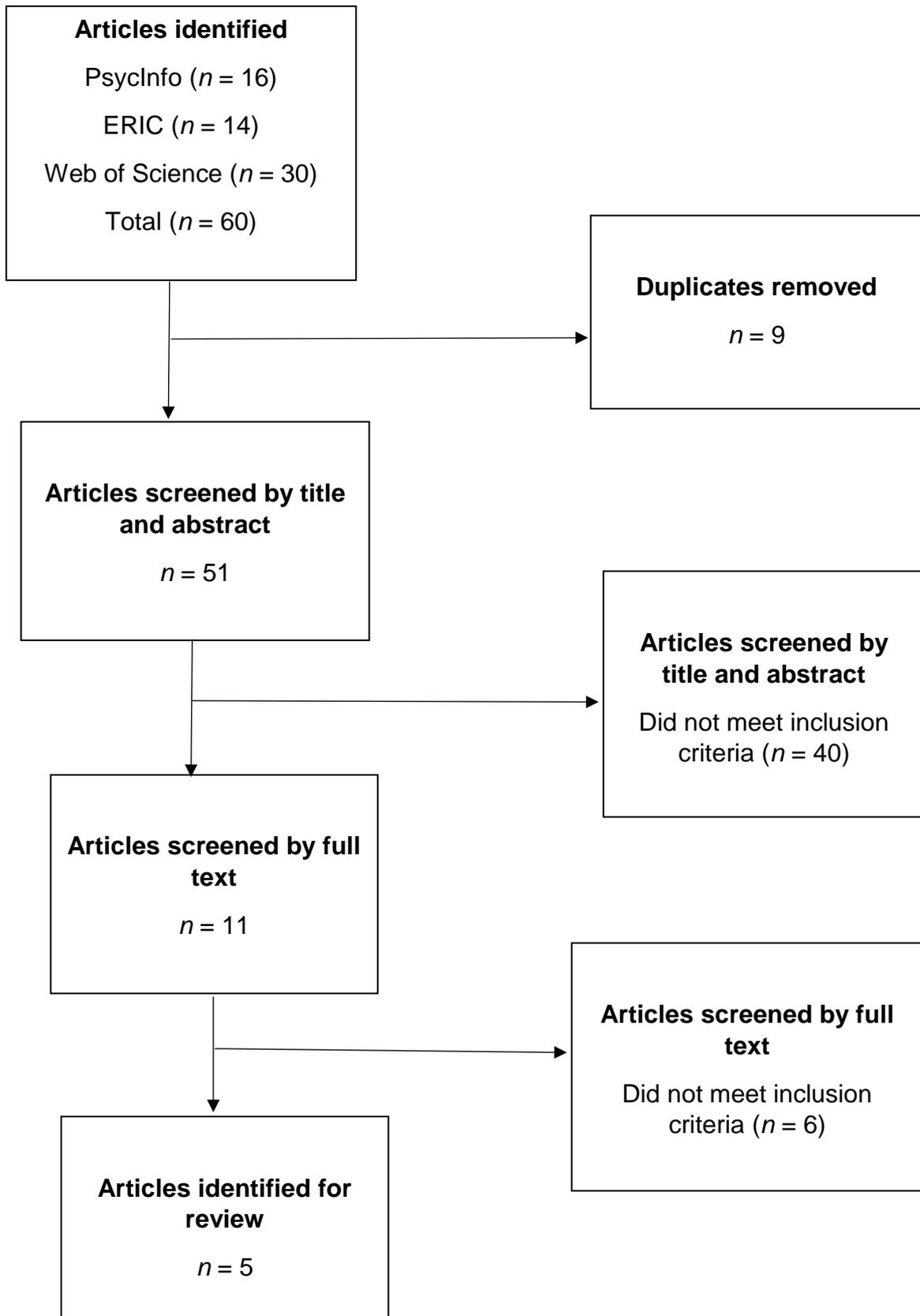


Table 2

Full List of Studies in this Review

Full Study Reference	
1	Lenihan, D., McCobb, E., Diurba, A., Linder, D., & Freeman, L. (2016). Measuring the Effects of Reading Assistance Dogs on Reading Ability and Attitudes in Elementary Schoolchildren. <i>Journal of Research in Childhood Education</i> , 30(2), 252-259.
2	Linder, D. E., Mueller, M. K., Gibbs, D. M., Alper, J. A., & Freeman, L. M. (2018). Effects of an animal-assisted intervention on reading skills and attitudes in second grade students. <i>Early Childhood Education Journal</i> , 46(3), 323-329.
3	Levinson, E. M., Vogt, M., Barker, W. F., Jalongo, M. R., & Van Zandt, P. (2017). Effects of reading with adult tutor/therapy dog teams on elementary students' reading achievement and attitudes. <i>Society & Animals</i> , 25(1), 38-56.
4	Uccheddu, S., Albertini, M., Pierantoni, L., Fantino, S., & Pirrone, F. (2019). The Impacts of a Reading-to-Dog Programme on Attending and Reading of Nine Children with Autism Spectrum Disorders. <i>Animals</i> , 9(8), 491.
5	Le Roux, M. C., Swartz, L., & Swart, E. (2014). The effect of an animal-assisted reading program on the reading rate, accuracy and comprehension of grade 3 students: A randomized control study. <i>Child & Youth Care Forum</i> , 43(6), 655-673.

3.3 Weight of Evidence

Using the Gough (2007) Weight of Evidence (WoE) Framework the five studies in this review were evaluated for their methodological quality and relevance, and moreover to consider the extent to which they answer the review question. Harden and Gough (2012) describe three key areas of weighing the quality and relevance of studies; WoE A methodological quality, WoE B methodological relevance, WoE C topic relevance to the review question. These WoE contribute to an overall WoE D (see Table 3 below).

WoE A was assessed using the Gersten, Fuchs, Compton, Coyne, Greenwood & Innocenti (2005) coding protocol, designed for use with group experimental studies (see Appendix C). This protocol was adapted as appropriate to critically appraise the research within this review (see Appendix F for details). For WoE B the Petticrew and Roberts (2003) typology was employed due to the nature of the review question (see Appendix D), and for WoE C a user-defined criteria was produced (see Appendix E) to examine the topic relevance to the review question.

Table 3

Weight of Evidence for Studies

Studies	WoE A	WoE B	WoE C	WoE D
Lenihan et al. (2016)	1	3	2	2 (Medium)
Linder et al. (2018)	1	3	2	2 (Medium)
Levinson et al. (2017)	3	3	2	2.67 (High)
Uccheddu et al. (2019)	2	3	2.25	2.42 (Medium)
Le Roux et al. (2014)	3	3	2.5	2.83 (High)

Low 1.4 or less, Medium 1.5-2.4, High 2.5 or above

Participants

A total of 199 participants were recruited into the five studies in this review, the studies ranged from nine participants in the smallest study (Uccheddu, Albertini, Pierantoni, Fantino & Pirrone, 2019) to 102 in the largest (Le Roux et al., 2014). These participants ranged in age from six to thirteen years of age, with three studies having a wide range of ages (Le Roux et al., 2014; Levinson, Vogt, Barker, Jalongo & Van Zandt, 2017; Uccheddu et al., 2019), and two studies solely looking at one year group (Lenihan, McCobb, Diurba, Linder & Freeman, 2016; Linder, Mueller, Gibbs, Alper & Freeman, 2018). These studies characterise a good age range found within primary schools and as a result provide a representative sample.

When exploring WoE C, one important criteria for consideration was the location of participants, with the United Kingdom (UK) having the most relevance to the review, followed by OECD countries and finally non-OECD countries. The studies ranged in location from the United States of America (USA), Italy and South Africa. Le Roux et al. (2014) scored lowest for this criterion in WoE C due to being in South Africa, which is the least likely to be comparable to the UK education system. The remaining four studies were from OECD countries (Lenihan et al., 2016; Levinson et al., 2017; Linder et al., 2018; Uccheddu et al., 2019) which are likely to be a closer match to the UK education system due to their social and economic similarities, and therefore they are more likely to be representative of schools in the UK. None of the studies scored the highest rating for location in WoE C, with none of the studies conducted in the UK.

Design

All five studies were categorised as randomised control trials (RCTs), resulting in receiving the highest rating for WoE B. RCTs are considered to be the highest quality methodological design, seen as 'gold standard' in relation to efficacy research (Petticrew & Roberts, 2003), and RCTs were also the most appropriate for the review question. Despite the random allocation of participants to groups in RCTs, these can vary tremendously in the procedure of the experimental and control groups. For example, the control group serves as a comparison to the experimental condition however if the control do not partake in an additional activity or treatment to usual (no treatment control), it is unclear if the change occurred due to the intervention itself or another factor such as the attention received. In contrast, by participants actively partaking in an alternative activity or treatment (treatment control) the effect of the experimental group (intervention) can be measured more accurately. The experimental groups were similar in content, however the control groups varied greatly

and this was reflected in the WoE C. Linder et al. (2018) scored the lowest for WoE C control group criteria due to the participants in the control group not receiving an alternative treatment but continued their normal school and home-based school activities- in this case reading at home (no treatment control). In comparison, a number of studies used a treatment control, and received a medium rating for WoE C control group criteria, where the participants read to an adult (Lenihan et al., 2016; Uccheddu et al., 2019) or a peer (Levinson et al., 2017). Le Roux et al. (2014) received the highest rating for WoE C control group criteria, having multiple control groups, both encompassing a no treatment control where participants continued with their normal school activities and two treatment controls with participants reading to an adult or teddy bear.

Whilst in most studies participants only took part in the experimental or control condition (Lenihan et al., 2016; Le Roux et al., 2014; Linder et al., 2018; Uccheddu et al., 2019), one of the studies in this review involved participants taking part in both conditions, using a repeated measures design (Levinson et al., 2017). This makes the effect of the intervention harder to interpret, due to half the participants having already experienced reading to a dog for a number of weeks before then taking part in the control condition where they read to a peer.

Intervention

The intervention consisted of the participants reading to a therapy dog, and as a result these did not differ in content due to the nature of the intervention. Where the interventions did vary was in the duration and number of sessions. Overall the total duration of the intervention was different across the studies, with the longest being 5 hours across 10 weeks (Uccheddu et al., 2019), followed by 3 hours reading (6 hours

total with the dog) over 6 weeks (Linder et al., 2018), and a number of the studies were approximately 2.5 hours over 5 weeks (Lenihan et al., 2016; Le Roux et al., 2014; Levinson et al., 2017). As seen above the number of weeks ranged from five to ten and the amount of time and the number of session in the intervention per week was different ranging from 15 minutes per session (Levinson et al., 2017; Le Roux et al. 2014) to 30 minutes per session (Lenihan et al., 2016; Linder et al., 2018; Uccheddu et al., 2019) as well as once (Le Roux et al. 2014; Lenihan et al., 2016; Linder et al., 2018; Uccheddu et al., 2019) or twice (Levinson et al., 2017) per week. Consequently, the intensity of the intervention varied greatly.

Measures

All studies measured reading accuracy, either with curriculum-based (Lenihan et al., 2016; Levinson et al., 2017; Linder et al., 2018) or standardised reading assessments (Le Roux et al., 2014; Uccheddu et al., 2019). Additional measures of reading ability were examined in the studies, including comprehension by some (Le Roux et al., 2014; Linder et al., 2018; Uccheddu et al., 2019) but not all studies, as a result the review focuses on reading accuracy.

Outcome Measures and Effect Sizes

Outcome measures are reported for reading accuracy for all five studies in the review. These were measured consistently despite the use of different instruments. Some additional measures of reading ability were reported for some studies however there was not consistent measures resulting in these not being reported as part of this review. The results presented focus on differences in control and intervention groups on reading ability to assess the impact of therapy dogs (see Table 4). The effect sizes

reported are Cohen's *d*, and where effect sizes were not available they were calculated permitting that sufficient information was present to do this.

Lenihan et al. (2016) had a medium effect (as displayed in Table 4) despite a small sample size. The results indicate that there is a strong relationship between the variables, thus given a larger sample size the findings may have had a larger effect size. Le Roux et al. (2014) reported a small effect size, despite the impact of a larger sample size. Le Roux et al. (2014) had the largest sample in the studies in this review, and also received the highest overall rating for methodological rigour and relevance on WoE D. Research by Levinson et al. (2017) found a range of effect sizes across a range of year groups from small effects for the oldest children measured in 5th Grade, medium for both 3rd and 4th Grades and large for the youngest children in 2nd Grade. These samples per year group were comparatively some of the smallest in comparison to other studies in this review. This study demonstrates the impact of therapy dogs specifically with a range of ages, and interestingly the effects on reading ability are larger for the youngest children and the effect sizes get smaller the older the children become. This has implications when considering at what age this intervention would be most effective.

Uccheddu et al. (2019) reported no effect between the variables, however this was the smallest sample with just nine participants and therefore it is likely any real effects would not be detected with such small representation of the population. Similarly, Linder et al. (2018) found no effect of intervention (see Table 4). This study received the lowest overall WoE D, with a low rating overall on WoE A. There was a lack of clarity in the procedures in how the intervention was approached in terms of set up for reading and what the participants were told. In addition the control group

was a non-treatment control with little known about the allocated reading time the children had or what was done during this time.

Table 4

Effect Sizes and Overall Quality Indicators

Study	Measures of Reading Accuracy	Number of Participants	Sample	Effect Size (Cohen's d)	Descriptor	WoE D
Lenihan et al. (2016)	Curriculum-based measures	15	2 nd Grade	.79	Medium	2 (Medium)
Linder et al. (2018)	DIEBELS (standard class assessment)	28	2 nd Grade	-.11	No Effect	2 (Medium)
Levinson et al. (2017)	Curriculum-based measures	11	2 nd Grade	.93	Large	
		13	3 rd Grade	-.68	Medium	2.67
		10	4 th Grade	.59	Medium	(High)
		11	5 th Grade	.28	Small	
Uccheddu et al. (2019)	Cornoldi Reading Test	9	6-11 years of age	.00	No Effect	2.42 (Medium)
Le Roux et al. (2014)	Neale Analysis of Reading Ability	102	7-13 years of age	.29	Small	2.83 (High)

*Reading accuracy was the only common measure of reading ability across studies – the review question focuses on this outcome measure
Cohen's d effect size: small = .20, medium = .50, large = .80*

4. Conclusions and Recommendations

This literature review explored to what extent one-to-one therapy dog interventions are effective at enhancing reading abilities of primary school children. The research findings for this intervention are mixed, with some large (Levinson et al., 2017), medium (Lenihan et al., 2016; Levinson et al., 2017), and small effect sizes (Le Roux et al., 2014) as well as no effects found (Linder et al. 2018; Uccheddu et al., 2019). Whilst there were a number of effects shown, due to the variation in intervention details, small sample sizes and methodological quality of each study it is difficult to make conclusive statements regarding the overall effectiveness. Larger samples would give more representative findings of a whole population.

Of the studies that found no effect, Uccheddu et al. (2019) had the smallest sample size overall and it is likely that the study was underpowered with such a small sample. Linder et al. (2018) also demonstrated no effect, however this study did not have a treatment control group. Instead the control group was described as continuing with everyday lessons and a comparison was made to the after school reading homework the children were set being similar to the additional reading time the intervention group experienced. Despite this, there was no understanding of the reading experience at home, or in fact if the controls felt more comfortable reading at home, or were also reading to a dog. This study received the lowest overall WoE D, with a low rating overall on WoE A. There was a lack of clarity in the procedures in how the intervention was approached in terms of the set up for reading and what the participants were told. This may have also led to less clarity for those running the intervention if these details had not been specified.

The research in this review involved the same intervention, however therapy dog interventions are not prescribed to the finer details. Consequently, while many reading programmes will be defined to run for a set number of weeks and session lengths, the specifics of therapy dog interventions are more indistinct. This has resulted in programmes that varied in the overall duration of intervention, with some studies having double the amount of intervention time than others, as well as for double the number of sessions and overall weeks of intervention. The intensity and the repetition of reading sessions may have had an unknown effect on reading ability. Similarly, most programmes were only five weeks with short durations of sessions. To consider the efficacy of such a programme these factors need to be considered in future research.

Despite post-test follow up in reading abilities, only one study (Le Roux et al., 2014) examined the longer term impact of the intervention on reading ability. With reading development being a gradual process it is possible that a lack of long-term follow up from the majority of the studies could have led to a missed effect of the intervention. Indeed there may have been an effect on reading but it would have taken longer to find observable measureable differences.

Levinson et al. (2017) was the only study to evaluate the impact of a therapy dog on reading abilities in different age ranges. This provided insightful evidence as to when it is most successful to intervene and when the effectiveness of these interventions become less so, although still effective overall. This has implications as to recognising when reading interventions may be most effective and should be considered by future researchers as well as with government initiatives.

Recommendations for future research would be to replicate the studies in this review, giving more standardised instructions to all participants, and using a substantially

larger sample size to ensure better generalisation of the findings. Interestingly, whilst a number of control groups were used, some of which were treatment controls, none of the comparison groups were for alternative programmes. It may be a consideration for future research to see how effective therapy dogs are compared to more accessible and less costly programmes to facilitate reading abilities. Equally, consideration of how long the sessions were and how many weeks would aid understanding of what would make a therapy dog intervention more effective.

At the present time, it would be difficult for Educational Psychologists to recommend the use of a therapy dog to facilitate reading abilities in primary school children. However, there is indicative evidence that there may be positive effects, and Educational Psychologists could be involved in conducting future research in the area. Despite this, therapy dogs are used for a range of purposes in schools, and the impact a therapy dog can have on other areas of development have not been considered in this review.

5. References

- Allen, G. (2011). Early intervention: the next steps. HM Government
- Bark and Read. (n.d.). Retrieved from <https://www.thekennelclub.org.uk/barkandread>
- Connell, C. G., Tepper, D. L., Landry, O., & Bennett, P. C. (2019). Dogs in Schools: The Impact of Specific Human–Dog Interactions on Reading Ability in Children Aged 6 to 8 Years. *Anthrozoös*, 32(3), 347-360.
- Department for Education and Department of Health. (2015). Special educational needs and disability code of practice: 0 to 25 years.
- Department for Education. (2019). National curriculum assessments: key stage 2, 2019 (revised).
- Department of Children, Schools and Families (2003). Every Child Matters.
- Friedmann, E., Katcher, A., Thomas, S., Lynch, J., & Messent, P. (1983). Social interaction and blood pressure: Influence of animal companions. *The Journal of Nervous and Mental Disease*, 171(8), 461–465.
- Friedmann, E., Thomas, S. A., & Eddy, T. J. (2000). Companion animals and human health: Physical and cardiovascular influences. *Companion animals and us: Exploring the relationships between people and pets*, 125-142.
- Friesen, L. (2010). Exploring animal-assisted programs with children in school and therapeutic contexts. *Early childhood education journal*, 37(4), 261-267.
- Geist, T. S. (2011). Conceptual framework for animal assisted therapy. *Child and Adolescent Social Work Journal*, 28(3), 243.
- Gersten, R., Puchs, L. S., Compton, D., Greenwood, C., & Innocenti, M. S. (2005). Quality Indicators for Group Experimental and Quasi-Experimental Research in Special Education. *Exceptional Children*, 71(2), 149–164.
- <https://doi.org/10.1177/001440290507100202>

Gough, D. (2007). Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Research papers in education*, 22(2), 213-228.

Harden, A., & Gough, D. (2012). Quality and relevance appraisal. Sage Publications.

Hughes, K. (2002). See Spot read. *Public Libraries*, 41(6), 328–330.

Jalongo, M. R., Astorino, T., & Bomboy, N. (2004). Canine visitors: The influence of therapy dogs on young children's learning and well-being in classrooms and hospitals. *Early Childhood Education Journal*, 32(1), 9-16.

Johnson, P., & Kossykh, Y. (2008). *Early Years, Life Chances and Equality: A Literature Review*. Manchester: Equality and Human Rights Commission.

Kirnan, J., Siminerio, S., & Wong, Z. (2016). The impact of a therapy dog program on children's reading skills and attitudes toward reading. *Early Childhood Education Journal*, 44(6), 637-651.

Le Roux, M. C., Swartz, L., & Swart, E. (2014). The effect of an animal-assisted reading program on the reading rate, accuracy and comprehension of grade 3 students: A randomized control study. *Child & Youth Care Forum*, 43(6), 655-673.

Lenihan, D., McCobb, E., Diurba, A., Linder, D., & Freeman, L. (2016). Measuring the Effects of Reading Assistance Dogs on Reading Ability and Attitudes in Elementary Schoolchildren. *Journal of Research in Childhood Education*, 30(2), 252-259.

Levinson, E. M., Vogt, M., Barker, W. F., Jalongo, M. R., & Van Zandt, P. (2017). Effects of reading with adult tutor/therapy dog teams on elementary students' reading achievement and attitudes. *Society & Animals*, 25(1), 38-56.

- Linder, D. E., Mueller, M. K., Gibbs, D. M., Alper, J. A., & Freeman, L. M. (2018). Effects of an animal-assisted intervention on reading skills and attitudes in second grade students. *Early Childhood Education Journal*, 46(3), 323-329.
- Parshall, D. P. (2003). Research and reflection: Animal-assisted therapy in mental health settings. *Counseling and Values*, 48(1), 47-56.
- Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology and Community Health*, 57, 527–529.
- Reichert, E. (1998). Individual counseling for sexually abused children: A role for animals and storytelling. *Child and Adolescent Social Work Journal*, 15(3), 177-185.
- Rose, J. (2006). Independent review of the teaching of early reading. *Department for Education and Skills*.
- Save the Children (2015). Ready to Read. London, UK.
- Schretzmayer, L., Kotrschal, K., & Beetz, A. (2017). Minor immediate effects of a Dog on children's reading Performance and Physiology. *Frontiers in veterinary science*, 4, 90.
- Uccheddu, S., Albertini, M., Pierantoni, L., Fantino, S., & Pirrone, F. (2019). The Impacts of a Reading-to-Dog Programme on Attending and Reading of Nine Children with Autism Spectrum Disorders. *Animals*, 9(8), 491.
- Wohlfarth, R., Mutschler, B., Beetz, A., & Schleider, K. (2014). An investigation into the efficacy of therapy dogs on reading performance in 6–7 year old children. *Human-Animal Interaction Bulletin*, 2(2), 60-73.

6. Appendices

Appendix A: List of Excluded Studies

Table 6.1

List of Excluded Studies at Full Text Screening

Reference	Reason for Exclusion
Moore, E. G., Hlava, T., Garcia, S. L., & Brem, S. L. (2013). Ever try teaching a dog to read? Implicit theories of reading ability. <i>Anthrozoös</i> , 26(3), 381-393.	Exclusion Criteria 5
Kirnan, J., Siminerio, S., & Wong, Z. (2016). The impact of a therapy dog program on children's reading skills and attitudes toward reading. <i>Early Childhood Education Journal</i> , 44(6), 637-651.	Exclusion Criteria 3
Kirnan, J., Ventresco, N. E., & Gardner, T. (2018). The impact of a therapy dog program on children's reading: Follow-up and extension to ELL students. <i>Early Childhood Education Journal</i> , 46(1), 103-116.	Exclusion Criteria 3
Kropp, J. J., & Shupp, M. M. (2017). Review of the Research: Are Therapy Dogs in Classrooms Beneficial? <i>Forum on Public Policy Online</i> . Oxford Round Table. 406 West Florida Avenue, Urbana, IL 61801.	Exclusion Criteria 4
Connell, C. G., Tepper, D. L., Landry, O., & Bennett, P. C. (2019). Dogs in Schools: The Impact of Specific Human–Dog Interactions on Reading Ability in Children Aged 6 to 8 Years. <i>Anthrozoös</i> , 32(3), 347-360.	Exclusion Criteria 3

Reference	Reason for Exclusion
Fung, S. C. (2019). Effect of a canine-assisted read aloud intervention on reading ability and physiological response: A pilot study. <i>Animals</i> , 9(8), 474.	Exclusion Criteria 2

Appendix B: Mapping the Field Table

Table 6.2

A Summary of Key Features for Studies in Review

No	Author	Participant s	Country	Study Design	Intervention	Measures	Primary Outcomes
1	Lenihan et al. (2016)	15 x 2 nd Grade (18 at the start)	USA	Independent samples. Random allocation to intervention and treatment control (dog and non-dog) groups	5 weeks once per week for 30 minutes	Curriculum-based measures	CBM reading (3 texts for 1 min) number of correct words per minute.
2	Linder et al. (2018)	28 x 2 nd Grade	USA	Independent samples. Random selection for those eligible. Random allocation to intervention and no treatment control (dog and non-dog) groups	6 weeks once per week for 1 hour (30 minutes reading)	DIEBELS (standard class assessment)	Number of words correct per minute, also looks at errors and (1 minute) comprehend a written text

No	Author	Participant s	Country	Study Design	Intervention	Measures	Primary Outcomes
3	Levinson et al. (2017)	45 x 2 nd -5 th Grades	USA	Repeated measures. Random stratified. Once selected randomly allocated (stratified on sex and grade) to intervention and treatment control (read to peer).	5 weeks (twice per week for 15 minutes)	Curriculum-based measures	Oral fluency (pre) mid (week 5) post (after week 10)
4	Uccheddu et al. (2019)	9 x 6-11 years (with Autism)	Italy	Random allocation based on demographics and diagnosis severity. Intervention and treatment control	10 sessions (once per week for 30 minutes)	Cornoldi reading test, test of reading comprehension, metaphonological competence	Reading fluency (speed and accuracy), reading comprehension, phonological awareness
5	Le Roux et al. (2014)	102 x 7-13 years (3 rd grade)	South Africa	3 intervention groups (dog, adult, teddy bear) and no treatment control.	10 sessions (once per	ESSI Reading and spelling test	Reading rate, accuracy and comprehension.

No	Author	Participant s	Country	Study Design	Intervention	Measures	Primary Outcomes
					week for 15- 20 minutes)		Measured 3 times (pre, post and delayed)

Appendix C: WoE A

Using the adapted Gersten et al. (2005) coding protocol the following ratings were found for WoE A (see Table 6.3). This protocol was edited to account for a typical rather than atypical population. Please see Appendix F for the Gersten et al. (2005) coding protocol and for the full details of this adaptation.

Table 6.3

Summary of Ratings for WoE A

Studies	Number of Essential Indictors	Number of Desirable Indictors	Overall WoE A
Lenihan et al. (2016)	8	3	1
Linder et al. (2018)	8	3	1
Levinson et al. (2017)	10	4	3
Uccheddu et al. (2019)	9	3	2
Le Roux et al. (2014)	10	4	3

Note: Overall WOE A (3 = High Quality; 2 = Acceptable Quality; <2 = Poor Quality)

Appendix D: WoE B

Using the Petticrew and Roberts (2003) guidelines for quality and rigour of research evidence, WoE B was defined as seen in Table 6.4 and coded (see Table 6.5). In this review using Petticrew and Roberts (2003) was considered appropriate to define WoE B, with the most effective way of measuring impact being experimental design and using an RCT being the best method to do this.

Table 6.4

WoE B Coding Criteria

Criteria	3 – High	2 – Medium	1 – Low
Study Design	Randomised control trials	Quasi-experimental, single-case study design or cohort studies	Non-experimental, case study, qualitative or surveys

Table 6.5

WoE B Ratings

Studies	WoE B
Lenihan et al. (2016)	3
Linder et al. (2018)	3
Levinson et al. (2017)	3
Uccheddu et al. (2019)	3
Le Roux et al. (2014)	3

Appendix E: WoE C

Criteria were defined to be appropriate for addressing the research question in this review (see Table 6.6). The studies in this review were then weighted for each of the criteria to produce an overall WoE C (Table 6.7).

Table 6.6

WoE C Criteria

Criteria	3 – High	2 – Medium	1 – Low
A. Location	UK	Intervention is conducted in a country that is economically similar- OECD	Intervention is conducted in a country that is not economically similar- non-OECD
B. Control Group	There are two or more control groups – attends session outside the normal classroom routine	There is one control group – attends session outside the normal classroom routine	Do not attend an additional session outside of the normal classroom routine
C. Depth of outcome measures	Multiple measures relating to reading	Single measure of reading	No measure of reading
D. Follow Up	Includes a follow up some time after post-test to consider impact of intervention on reading ability over time	Includes post-test findings to evaluate impact of the intervention	No follow up data

Table 6.7

WoE C Overall Ratings

Studies	Criteria A	Criteria B	Criteria C	Criteria D	WoE C
Lenihan et al. (2016)	2	2	2	2	2
Linder et al. (2018)	2	1	3	2	2
Levinson et al. (2017)	2	2	2	2	2
Uccheddu et al. (2019)	2	2	3	2	2.25
Le Roux et al. (2014)	1	3	3	3	2.5

Appendix F: Coding Protocol

Gersten et al's (2005) coding protocol was used and adapted to suit the range of studies in this review. Specifically, the participants were from a typical population in this review rather than from a population with specific disabilities or difficulties. In response to this, Question 1 of the Gersten et al. (2005) protocol was edited to become relevant to the target population (see below).

Coding Protocol: Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C, & Innocenti, M. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, 71,149-164.

Study:

Author:

Essential Quality Indicators - Quality indicators for describing participants

~~Was sufficient information provided to determine/confirm whether the participants demonstrated the disability(ies) or difficulties presented?~~

Yes

No

N/A

Unknown/Unable to Code

Was sufficient information provided on participants (Age, Ethnicity, Social Economic Status, difficulties/disabilities)?

Yes

No

N/A

Unknown/Unable to Code

Did the study use appropriate participants for the research question?

Yes

- No
- N/A
- Unknown/Unable to Code

Were appropriate procedures used to increase the likelihood that relevant characteristics of participants in the sample were comparable across conditions?

- Yes
- No
- N/A
- Unknown/Unable to Code

Was a sufficiently large sample size used?

- Yes
- No
- N/A
- Unknown/Unable to Code

Total size of sample (start of study): _____

Were participants randomly assigned to conditions?

- Yes
- No
- N/A
- Unknown/Unable to Code

Did the research include a control group?

- Yes
- No
- N/A
- Unknown/Unable to Code

Total size of control group (start of study): _____

Was sufficient information given characterizing the interventionists or teachers provided? Did it indicate whether they were comparable across conditions?

- Yes
- No
- N/A
- Unknown/Unable to Code

Essential Quality Indicators - Quality indicators for Implementation of the Intervention and Description of Comparison Conditions

Was the intervention clearly described and specified?

- Yes
- No
- N/A
- Unknown/Unable to Code

Did the research specify the length of intervention?

- Yes
- No
- N/A
- Unknown/Unable to Code

Was the fidelity of implementation described and assessed?

- Yes
- No
- N/A
- Unknown/Unable to Code

Was the nature of services provided in comparison conditions described?

- Yes
- No

- N/A
- Unknown/Unable to Code

Essential Quality Indicators – Quality Indicators for Outcome Measures

Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention and measures of generalised performance?

- Yes
- No
- N/A
- Unknown/Unable to Code

Were outcomes for capturing the intervention’s effect measured at the appropriate times?

- Yes
- No
- N/A
- Unknown/Unable to Code

Did the research use well-known or standardized measures?

- Yes
- No
- N/A
- Unknown/Unable to Code

Essential Quality Indicators – Quality Indicators for Data Analysis

Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the limit of analysis in the study?

- Yes
- No
- N/A
- Unknown/Unable to Code

Did the research report include not only inferential statistics but also effect size calculations?

- Yes
- No
- N/A
- Unknown/Unable to Code

Essential Quality Indicators Total Score:

Desirable Quality Indicators

Was data available on attrition rates among intervention samples? Was severe overall attrition documented? If so, is attrition comparable across samples? Is overall attrition less than 30%?

- Yes
- No
- N/A
- Unknown/Unable to Code

Did the study provide not only internal consistency reliability but also test-retest reliability and interrater reliability (when appropriate) for outcome measures? Were data collectors and/or scorers blind to study conditions and equally (un)familiar to examinees across study conditions?

- Yes
- No
- N/A
- Unknown/Unable to Code

Were outcomes for capturing the intervention's effect measured beyond an immediate post-test?

- Yes
- No
- N/A
- Unknown/Unable to Code

Was evidence of the criterion-related validity and construct validity of the measures provided?

- Yes
- No
- N/A
- Unknown/Unable to Code

Did the research team assess not only surface features of fidelity implementation (e.g. number of minutes allocated to the intervention or teacher/interventionist following procedures specified), but also examine quality of implementation?

- Yes
- No
- N/A
- Unknown/Unable to Code

Was any documentation of the nature of instruction or series provided in comparison conditions?

- Yes
- No
- N/A
- Unknown/Unable to Code

Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?

- Yes
- No
- N/A
- Unknown/Unable to Code

Were results presented in a clear, coherent fashion?

- Yes
- No

N/A

Unknown/Unable to Code

Essential Quality Indicators Total Score:

	Total	Score
Essential Quality Indicators Total of >9 = Score 1 Total of <9 = Score 0		
Desirable Quality Indicators Total of ≥4 = Score 2 Total of <4 = Score 1 Total 0 = Score 0		
Total Score (3 = High Quality; 2 = Acceptable Quality; <2 = Poor Quality)		

Study Rating:

Study: Lenihan, D., McCobb, E., Diurba, A., Linder, D., & Freeman, L. (2016). Measuring the Effects of Reading Assistance Dogs on Reading Ability and Attitudes in Elementary Schoolchildren. *Journal of Research in Childhood Education*, 30(2), 252-259.

Essential Quality Indicators - Quality indicators for describing participants

Was sufficient information provided to determine/confirm participants demographics?

Yes

No

N/A

Unknown/Unable to Code

Were appropriate procedures used to increase the likelihood that relevant characteristics of participants in the sample were comparable across conditions?

Yes

No

N/A

Unknown/Unable to Code

Was sufficient information given characterizing the interventionists or teachers provided? Did it indicate whether they were comparable across conditions?

Yes

No

N/A

Unknown/Unable to Code

Essential Quality Indicators - Quality indicators for Implementation of the Intervention and Description of Comparison Conditions

Was the intervention clearly described and specified?

Yes

No

N/A

Unknown/Unable to Code

Was the fidelity of implementation described and assessed?

Yes

No – described what they did but no standard way of doing in and not assessed

N/A

Unknown/Unable to Code

Was the nature of services provided in comparison conditions described?

Yes – read to human volunteer

No

N/A

Unknown/Unable to Code

Essential Quality Indicators – Quality Indicators for Outcome Measures

Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention and measures of generalised performance?

Yes

No

N/A

Unknown/Unable to Code

Were outcomes for capturing the intervention's effect measured at the appropriate times?

Yes – before and after

No

N/A

Unknown/Unable to Code

Essential Quality Indicators – Quality Indicators for Data Analysis

Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the limit of analysis in the study?

Yes

No

N/A

Unknown/Unable to Code

Did the research report include not only inferential statistics but also effect size calculations?

Yes

No

N/A

Unknown/Unable to Code

Essential Quality Indicators Total Score: 8

Desirable Quality Indicators

Was data available on attrition rates among intervention samples? Was severe overall attrition documented? If so, is attrition comparable across samples? Is overall attrition less than 30%?

Yes

No

N/A

Unknown/Unable to Code

Did the study provide not only internal consistency reliability but also test-retest reliability and interrater reliability (when appropriate) for outcome measures? Were data collectors and/or scorers blind to study conditions and equally (un)familiar to examinees across study conditions?

Yes

No – not blind? Does not specify

N/A

Unknown/Unable to Code

Were outcomes for capturing the intervention's effect measured beyond an immediate post-test?

Yes

No

N/A

Unknown/Unable to Code

Was evidence of the criterion-related validity and construct validity of the measures provided?

Yes – justified why used and benefits, measures what intended

No

N/A

Unknown/Unable to Code

Did the research team assess not only surface features of fidelity implementation (e.g. number of minutes allocated to the intervention or teacher/interventionist following procedures specified), but also examine quality of implementation?

Yes

No

N/A

Unknown/Unable to Code

Was any documentation of the nature of instruction or series provided in comparison conditions?

Yes

No

N/A

Unknown/Unable to Code

Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?

Yes

No

N/A

Unknown/Unable to Code

Were results presented in a clear, coherent fashion?

Yes – no effect sizes, S.D

No

N/A

Unknown/Unable to Code

Essential Quality Indicators Total Score: 3

	Total	Score
Essential Quality Indicators Total of >9 = Score 1 Total of <9 = Score 0	8	0
Desirable Quality Indicators Total of ≥4 = Score 2 Total of <4 = Score 1 Total 0 = Score 0	3	1
Total Score (3 = High Quality; 2 = Acceptable Quality; <2 = Poor Quality)		1

Study Rating: Poor Quality