

This is a repository copy of How the type and context of children's storybook reading relate to select empathy skills: a meta-analysis.

White Rose Research Online URL for this paper: <u>https://eprints.whiterose.ac.uk/228388/</u>

Version: Published Version

# Article:

Ciesielska, M., Kucirkova, N. and Thomson, J. orcid.org/0000-0002-4293-4851 (2025) How the type and context of children's storybook reading relate to select empathy skills: a meta-analysis. Early Education and Development. ISSN 1040-9289

https://doi.org/10.1080/10409289.2025.2516989

# Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/

# Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk https://eprints.whiterose.ac.uk/





Early Education and Development

ISSN: 1040-9289 (Print) 1556-6935 (Online) Journal homepage: www.tandfonline.com/journals/heed20

# How the Type and Context of Children's Storybook Reading Relate to Select Empathy Skills: A Meta-Analysis

Marta Ciesielska, Natalia Kucirkova & Jenny Thomson

**To cite this article:** Marta Ciesielska, Natalia Kucirkova & Jenny Thomson (17 Jun 2025): How the Type and Context of Children's Storybook Reading Relate to Select Empathy Skills: A Meta-Analysis, Early Education and Development, DOI: <u>10.1080/10409289.2025.2516989</u>

To link to this article: https://doi.org/10.1080/10409289.2025.2516989

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.



0

View supplementary material 🕝

4	0

Published online: 17 Jun 2025.



Submit your article to this journal 🗹



View related articles 🗹



View Crossmark data 🗹



OPEN ACCESS Check for updates

# How the Type and Context of Children's Storybook Reading Relate to Select Empathy Skills: A Meta-Analysis

Marta Ciesielska<sup>a</sup>, Natalia Kucirkova<sup>b,c</sup>, and Jenny Thomson<sup>d</sup>

<sup>a</sup>Human Communication Sciences and National Institute of Teaching, University of Sheffield; <sup>b</sup>Læringsjmiljøsenteret, University of Stavanger; Faculty of Education, The Open University; <sup>d</sup>Human Communication Sciences, University of Sheffield

#### ABSTRACT

This systematic literature review and meta-analysis explored the impact of storybook reading on four empathy-related skills (theory of mind, prosocial behavior, perspective-taking and emotional understanding), in relation to storybook format (digital/paper), content and type of reading (adult's role: active/ passive/or adult not present; intervention location and setting; individualized/ shared). Included studies compared empathy between experimental (storybook reading) and control groups of N = 2,293 2- to 10-old children. We analyzed 21 eligible studies from 10 countries published between 1977 and 2022. The metaanalysis revealed an overall impact on empathy, but this effect only held significance for children's prosocial skills. Research findings: There was no effect of format, and insufficient information to examine the effect of storybook content. No significant differences were found between children reading on their own or with an adult in an active or passive role. However, children reading in an independent or one-to-one context uniquely predicted overall empathy scores, while reading in groups did not. Practice or Policy: Our findings nuance the universal claim that storybook reading promotes children's empathy with a specification that this may concern only children's prosocial behavior and be best promoted in an individualized context.

Empathy is a key social skill for navigating the world, especially for children growing up in today's globalized and often fragmented societies. Empathy has been defined in various ways in the literature and for the purpose of our study, we adopted the operational definition by Cikara and colleagues who define empathy as an individual's understanding of the emotions of another person and acting on this understanding, is not innate but nurtured through practice and dedicated resources (Cikara et al., 2011). The key activity presupposed to nurture empathy skills in early childhood is reading of narratives in a textual and pictorial format (referred to in early childhood literature as storybook reading), but there is no summative evidence to support this assumption.

A number of nationally and internationally funded interventions seek, and claim, to enhance children's empathy through storybook reading (e.g. The Empathy Lab:https://www.empathylab.uk/, Roots of Empathy: https://rootsofempathy.org). However, the overall evidence on the relationship between storybook reading and children's empathy skills is not clear, with several hypothesized pathways between various types of storybooks and empathy-related skills (Kucirkova, 2019). Our study aims to provide systematic and meta-analytical evidence on the specific relationships between types of storybooks and storybook reading with skills related to empathy in children at the early and advanced stage of reading.

CONTACT Natalia Kucirkova 🖾 Natalia.kucirkova@uis.no 💽 Læringsjmiljøsenteret, University of Stavanger, Hulda Gaborgs Hus, Stavanger 4320, Norway.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10409289.2025.2516989.

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

There are three issues that, from an educational point of view, need to be addressed in a summative manner. First, it is not clear what kind of empathy skills could be promoted by storybook reading. One of the difficulties in establishing the impact of storybook-reading on empathy with children is that the use of stories is often part of larger interventions, which makes it difficult to disentangle its separate effect from the intervention's overall impact on learners (Betawi, 2022); furthermore, storybook reading studies without baseline and control group assessments are common in the field of shared book reading, which complicates meta-analytical efforts (see Noble et al., 2019).

The second key issue to examine relates to the conceptual inconsistencies in how researchers define and study empathy. Researchers agree that empathy is critical for children's healthy peer relationships (Bjorkqvist et al., 2000) and advanced social skills (Silvera et al., 2001). However, as a popular term widely used in public discourse, empathy has been defined inconsistently in research studies depending upon the investigational focus, particularly in terms of its cognitive or emotional elements (Cuff et al., 2016) and developmental versus neuroscientific studies, see (Abramson et al., 2020). Different definitions reflect the richness of the field on this topic but it can also result in misinterpretation of outcomes attributed to empathy (Brown et al., 2012).

Third, it is unclear which types and kinds of reading sessions and which aspects of texts are most suitable for empathy development. Narrative fiction reading can take various forms (digital, print, hybrid books) in various contexts (at home, pre-school or on the move), with children reading on their own or with their family members and teachers. The differential impact of individual context, content and child characteristics is a leading question in children's reading studies (e.g. Neuman et al., 2021; Yang et al., 2022), but has been neglected in summative empathy examinations. It could be that the evidence base is stronger for some types of reading than for others.

Reading improves foundational literacy developmental skills (see Sénéchal, 2017, for an overview), and our question is *how* storybook reading relates to children's empathy skills. We aimed to determine the magnitude of the impact of storybook reading on the empathy skills in children aged 2–10 years, and delineate the types of books and reading that promote specific empathy skills in young children.

Our focus on this age group was informed by two rationale. First, from a language processing perspective, the ability to make meaning of texts and images is independent from the ability to process alphabetic scripts (Berg, 1998). Thus, while the processing of written language changes rapidly in the 2–10-year-old age group, the oral language processing of narrative is not radically different within this age group. We therefore decided to focus on this age group given the literature consensus that this age range covers the most active stage of reading development for children across most cultures and considers it a critical period for children's language, social and cognitive development (Kucirkova et al., 2017). Second, while we did not aim to examine any developmental pathways in children's empathy, we wanted to include a broader age range for an examination of possibly differential impacts of storybook reading types and contexts on different age groups. For this objective, the 2 to 10-old age group seemed appropriate given that children typically examined for book reading studies are exposed to different types of books and reading contexts as they transition from infant to toddler and toddler to kindergarten and early primary school age (Sulzby, 1985; Whitehurst & Lonigan, 1998).

#### **Delineations and classifications**

Although somewhat simplified, it is useful to acknowledge interdisciplinary approaches that inform the rich body of literature on children's storybook reading and empathy. Broadly speaking, qualitative studies and analyses of children's literature highlight the esthetic, stylistic and literary characteristics that might be implicated in the empathy benefits that children derive from storybook reading (e.g. Nikolajeva, 2009, 2012). Conversely, cognitive and neuroscience approaches to reading are often interested in how readers' characteristics relate to various empathy outcomes, such as theory of mind or prosocial behavior (Bruneau et al., 2015, 2017). In our quantitatively oriented review and meta-analysis, we were interested in aspects that interdisciplinary theorists (e.g. Kümmerling-Meibauer, 2014; Mar, 2011) agree on as vital elements in empathy-related storybook reading, namely the importance of narrative and fictional character of stories, child's characteristics (age), context of reading (adult presence and interactivity) and type of storybooks read. Empirical studies with adult readers (e.g (Kidd & Castano, 2013) suggest that it is reading of literary narrative fiction, (as opposed to nonfiction), that supports empathy (Gottschall, 2012; Keen, 2007; Mar, 2011). For adult readers, questions remain regarding the impact of texts' literariness e.g., stylistic differences between texts, narrativity: the degree to which a text displays a plot trajectory, including character development and fictionality, or the "reality status" of the content and empathy (see Koopman, 2018). Yet for children, the overall presence of a relationship between empathy and storybook reading is still to be established. Furthermore, while for adult readers there is meta-analytical evidence for the impact of reading fiction books on a specific empathy skill – theory of mind (Dodell-Feder & Tamir, 2018; Mumper & Gerrig, 2017), there is no such summative evidence for children. This is a significant research gap given the emphasis on reading in childhood settings (e.g., Mason, 1990) and the rapid development and potential diverse trajectories of empathy in childhood (McDonald & Messinger, 2011).

#### Empathy

We reviewed the definition of empathy in studies from various research fields, including psychology, education, social sciences and neurocognitive sciences, in order to arrive at an interdisciplinary and encompassing definition of the term. Empathy's mental and behavior components include both positive and negative aspects of empathy (cf. Bloom, 2017) and both affective and cognitive sides (cf Cox et al., 2012). Interdisciplinary literature agrees that while emotional response to another person's emotions is innate; cognitive empathy (aka social cognition) is nurtured through practice and dedicated resources (Saxe, 2006). As a cognitive capacity, empathy includes social understanding, which subsumes emotion recognition, emotion understanding and perspective taking (Luke & Banerjee, 2013). In cognitive neuroscience, there is a distinction between two main dimensions of empathy: cognitive and emotional empathy (Blair, 2005). Emotional empathy is "the cognitive and neural processes that produce a congruent emotion in the observer in response to others' directly perceived emotional displays or to descriptions of others' emotion-laden experiences" (Saxe, 2006, p. 235), while cognitive empathy relates to the conscious process of engaging with someone's emotions. The two-dimensional aspect of empathy, one conscious and one automatic emotional state of feeling for someone, has been followed in psychology studies since the 1980s (Davis, 1983) and consolidated with neurological findings and dual-process theories (Yu & Chou, 2018).

In educational and social studies, empathy is defined as a sub-skill of social cognition, together with Theory of Mind (ToM) and emotion understanding (Ornaghi et al., 2014). Social cognition is used as a general term to describe cognition involving others or "the know-how that allows us to sustain interactions, form relations, understand each other, and act together" (De Jaegher et al., 2010, p. 331). In both psychology and education studies, empathy tends to be defined as an individual mechanism with an interaction component in the form of prosocial behavior. The latter highlights the implications of empathy for morally desirable behaviors (such as truth-telling, helping or caring for one another). Some scholars conceptualize empathy as a precursor to prosocial behavior (e.g. Hoffman, 2000), while others describe empathy as part of prosocial behavior (Bierhoff, 2002). Hoffman (2000) proposed that empathy is an emotional factor that motivates prosocial behavior but whether one engages in prosocial behavior depends on the type of emotions they experience with guilt being closely associated with prosocial behavior (Vaish, Carpenter, & Tomasello, 2016).

Based on this literature, our definition of empathy subsumes four core skills that play a dominant role in interdisciplinary studies of empathy and children's storybook reading: theory of mind, prosocial behavior, emotional vocabulary and understanding, and perspective taking. These four skills have a relatively widely agreed definition in the field, as follows:

Theory of Mind (ToM) is "the ability to understand the mental state of other people which includes others' thinking, beliefs, desire and emotion" (Wulandini et al., 2017, p. 2) and is a crucial skill for school-aged children to enable successful collaborations and establishments of

peer relationships as it organizes emotions and moderates empathic responses (Kidd & Castano, 2013).

Prosocial behavior is understood as a selfless act of helping others without expectations to get something in return (Bar-Tal, Raviv, & Goldberg, 1982). It comprises two major components: helping intentions and helping behavior. The former relies largely on moral judgments, whereas the latter is a form of realization of helping intentions (Tangney et al., 2007).

Emotional vocabulary and understanding refer to social emotional assets that are part of interpersonal skills, and an individual's self-concept, and that can improve students' social emotional competencies and are often studied as part of the bullying prevention literature (e.g. Wang et al., 2015).

Finally, perspective-taking is an individual's ability to infer others' emotions and feelings and can be both of an affective and cognitive kind (Healey & Grossman, 2018).

With the four skills included, we see empathy as a composite term that refers to an individual's understanding of the emotions of another person and acting on this understanding.

#### Storybooks and storybook reading

Our interest is in the effects of reading storybooks on children's empathy. We focus on storybook reading because the activity can simulate social situations and introduce a mental distance between the "self" and the "other" and thus potentially promote empathy-related skills (Orvell et al., 2019). Similar to other interdisciplinary meta-analytical studies concerned with children's reading and development (e.g. Fikrat-Wevers et al., 2021; Takacs et al., 2015) we define storybooks as illustrated narrative fiction, containing written words and a plot: in any format (paper-based or digital with and without interactive and multimedia features); commercially available, or researcher developed. Storybooks could include visual and audio features and were read by the child independently or by an adult to the child.

#### Literature review

#### Empathy and storybook reading

There are several activities that are hypothesized to promote empathy, including role-play and virtual reality games, but storybooks offer uniquely low-risk, noninvasive options for directly adopting the role of an unknown character (Kucirkova, 2019). With their fictional characters and unfamiliar settings, storybooks nurture children's imagination (Nikolajeva, 2009), which is linked to the ability to self-distance and help them infer characters' thoughts and feelings (Kümmerling-Meibauer, 2014). Furthermore, reading books offers rich opportunities for discussing various socio-emotional situations with others, as books depict and deal with both real-life and imagined experiences, with animate and inanimate story characters and their feelings, thoughts, intentions, and desires (Mar et al., 2009).

The research recognition that reading is a leading context for exposing children to the feelings of others and practicing perspective-taking, has been translated into several intervention studies examining the relationship between reading and empathy (e.g., Gloor & Puhl, 2016; Herrera et al., 2018). Notably in cognitive and neuroscientific research with adults, social cognition was linked to literary fiction reading (e.g. Tamir et al., 2016), through the mechanism of simulating social worlds by repeatedly reading about characters' experiences (Oatley, 2016). Although the positive effects of short-term interventions (Kidd & Castano, 2013) failed to be replicated (Panero et al., 2016), there is meta-analytical evidence of a small, statistically significant improvement in *adult* readers' social-cognitive performance after reading fiction versus nonfiction (Dodell-Feder & Tamir, 2018).

That empathy can be nurtured and trained throughout the lifespan is a frequently confirmed research finding (Baron-Cohen, 2011), which furthers the rationale to use lifelong activities, such as storybook reading, as the primary empathy-building context at an early age. However, as Spinrad and Gal (2018) concluded, more experimental research is needed to establish the conditions and

mechanisms through which empathy-building occurs. Furthermore, as shown next, studies indicate the potential impact of storybook reading on diverse components of empathy.

#### Types of empathy skills promoted through storybook reading

Individual studies have shown positive relations for each of the possible manifestations of empathy, although the overall effect remains to be investigated. The most abundant studies are in the area of theory of mind (ToM) or cognitive empathy. For example, reading literary fiction, as opposed to popular fiction and nonfiction, enhanced ToM performance in children aged nine to ten (Wulandini et al., 2017). In relation to children's prosocial behavior, Salay (2018) examined empathy development during a four-week intervention with first-graders and through a qualitative analysis found that the students showed empathetic and non-empathetic behaviors toward the story characters and people in both real life and imaginary scenarios in their writing assignments following the intervention.

As for emotional vocabulary and understanding, Schapira and Aram (2020) measured the impact of children's shared book reading experience on children's socio-emotional competence, which is similar to ToM and perspective-taking in that it enables children to appreciate the needs of others, notice their feelings and respond appropriately with prosocial behavior (Garner & Estep, 2001); this skill is also fundamental to a child's academic achievement and school adjustment (Denham et al., 2012), as well as future ability to form social relationships (Izard et al., 2001).

Regarding the relationship between perspective-taking and storybook reading, Gil et al. (2014) documented higher perspective-taking skills in six-year-old children who participated in an augmented-reality storybook reading. Furthermore, a longitudinal study with parents and their three to 4-year-old children found that parent-child perspective-taking during comic book reading contributed to pre-schoolers' false belief attribution in the following year (Bernard & Deleau, 2007).

The overall picture that emerges from the literature is that storybook reading can impact various children's empathy skills, the magnitude of which remains to be established. We therefore hypothe-sized that storybook reading will be related to all four empathy skills to a different level of magnitude.

The next question that emerges is whether this is the case for all types of storybooks, or whether some storybooks are more suitable for the promotion of empathy than others.

#### Types of books and children's empathy skills

Storybooks can serve as a vehicle for moral development but the type of content and format of storybooks matters in how effective this development process is (Hoffman, 2000). Cassidy et al. (1998) analyzed the content of children's popular literature and found that in more than 75% of the books that parents reported to read to their 3–5-year olds, there were references to mental states, false-beliefs and irony. In an empirical study, Lee et al. (2014) examined children's truth-telling (understood in the study as a component of empathy) in relation to four classic stories – The Tortoise and The Hare, Pinocchio, The Boy Who Cried Wolf, George Washington and the Cherry Tree – and found that only the reading of the last title increased truth-telling in 3- to 7-year-olds.

Given that reading of high-quality texts reliably predicts children's story comprehension, it is likely to be related to theory of mind (Adrian et al., 2005). However, for this relationship, the format of storybook reading might play a particularly important role, given that some digital books, notably those with distracting interactive features, impede children's reading comprehension (Furenes et al., 2021). Furthermore, the ToM and reading comprehension relationship is influenced by the content of the storybooks. Du and Hao (2018) found that moral stories have a facilitatory effect on prosocial behavior of pre-schoolers but that only moral stories that highlighted negative emotions toward non helping behavior moderated children's actual helping behavior. Also highlighting an interactive effect between book type and emotional valence, Schapira et al. (2021) compared mother-child discourse in narrative and didactic books with either negative or positive emotional foci. Mothers used the didactic books for more detailed discussion about the "negative" emotions of anger and sadness, specifically. The authors interpreted the simpler structure of didactic books as allowing more open discussion, and speculated that mothers may have felt that negative emotions needed more explanation in terms of how to deal with them.

Based on this literature, we hypothesized that the type of book, in relation to both format and content, will differently impact the effects of storybook reading on children's empathy skills.

#### Types of storybook reading and children's empathy books

The nature of shared adult-child interactions during reading, in addition to, or without reading storybooks, contributes to the development of children's ToM (Ornaghi & Grazzani, 2013). Notably, the frequency of parental utterances concerning mental states correlates with their children's ToM performance (Ruffman et al., 2002).

This raises the conceptual question of whether it is storybook reading per se, or the conversation and interaction around storybook reading, that contributes to children's ToM. Literature in both directions exists: An experimental study comparing children's gains on false belief tasks found higher effects for children who read with parents at home than those that didn't (Adrian et al., 2005). Yet, in an intervention experiment with low-income preschoolers, empathy gains were noted after children engaged in storybook discussions with an experimenter (Tompkins, 2015). Interestingly, the experimental group did not perform better than control groups on measures of emotion understanding and social competence. Tompkins (2015) found that children who participated in the intervention outperformed their peers in both control groups (one control group read books with no discussion, the second control group only engaged in discussions with no reading) on ToM performance, thus indicating that it is the combination of reading and discussion that impacts children's ToM skills. Not all parents engage in language- and emotion-stimulating storybook conversations, however, as shown by Aram et al. (2017), who found little discussion around the book with low SES mothers. Interestingly, for these mothers, it was the conversations with their children after the reading session that were associated with children's higher socio-emotional understanding and prosocial behavior.

The role of an adult in reading needs to be distinguished not only in terms of presence/absence and the role of a parent or experimenter (researcher) but also in terms of interactivity, or how active or passive adults are during the reading activity with the child. Mar et al. (2010) proposed that unlike other parent-child activities, storybook reading enables active involvement of an adult in discussions regarding story characters, their feelings, intentions, and thoughts. This in turn allows the child to gradually better understand the mental states of others as well as of their own. In addition, storybook reading acts as a mental simulation that supports imagination development which in consequence predicts social competence in children (Goodman & Dent, 2019). The adults' active role seems to be a key component in these relationships.

In addition to the ways adults participate in the activity, storybook reading sessions can vary through being shared/joint with a parent in a parent-child dyad, one-to-one with an experimenter or by children reading in groups or alone. Experiments with online reading show that reading comprehension is higher in groups of children reading together than individually (Mitra & Crawley, 2014; Vega et al., 2020). Group reading in classrooms can be designed to be more coordinated with digital interfaces (Pearson et al., 2012), which increases the intervention possibility for children to engage in coordinated, empathy-boosting peer discussions. The book genre influences the extent to which parents engage in empathy-promoting talkfor example, narrative wordless picture books for toddlers stimulated more complex maternal talk than didactic books (Nyhout & O'Neill, 2013).

In an effort to consolidate this rich and often fragmented literature, we aimed to examine the relationships between storybook reading and empathy in a systematic way, guided by an interdisciplinary focus on the type of reading (group/individualized reading with/without parent) and reading interactivity operationalized here as adults' active or passive role. We expected that shared storybook reading sessions with high interactivity would have the strongest empathy effect.

# Study aims

Our aim was to establish the strength of different story books and reading types on different children's empathy skills. Understanding the link between storybooks, reading types, and empathy in children is crucial. As discussed in the literature review, empathy is a cornerstone of social interaction, allowing us to connect with others and build positive relationships. By pinpointing how specific stories and reading approaches influence empathy development, we can create targeted strategies to nurture this vital skill in children. This knowledge can empower educators, parents, and caregivers to foster a love of reading that goes beyond literacy, promoting emotional intelligence and well-adjusted young minds.

Based on our literature review, we developed a logic model that specifies the different moderators that play a role in how storybooks impact children's empathy skills. Our logic model is in Rehfuess et al. (2018) definition an a priori, process-oriented logic model that captures the complex relationship between moderators and outcomes in the proposed storybook–empathy relationship. The model (see Figure 1) depicts the key storybook characteristics as an index of reading interventions implicated in children's empathy skills. The model further depicts how different storybook characteristics and reading types can act as moderators or mediators that influence the outcome of empathy skills in children. Finally, the model depicts the diverse types of reading as moderators of the relationship between storybook reading and empathy skills.

Based on the model, we developed three main research questions (RQ), answering of which will inform the field by delineating several aspects related to the impact of empathy and storybook reading, including the types of skills. Books and reading supports involved:

RQ1: What is the impact of storybook reading on children's empathy-related skills and which empathy skills are most affected by storybook reading?

RQ2: Which storybooks promote empathy skills and does the format and/or content of storybooks play a role?



Figure 1. The Study's heuristic model.

RQ3: What kind of reading predicts empathy skills and does the presence of the adult, adults' active role during reading and setting of the reading session influence the results?

#### Methods

We systematically reviewed and meta-analyzed the available research concerned with the impact of storybook reading on four empathy-related skills: theory of mind, perspective-taking, emotional vocabulary and understanding, and prosocial behavior for children aged 2–10 years. We were also interested in identifying the content and format features that characterize empathy-promoting storybooks.

#### Inclusion/Exclusion criteria

Studies included in the meta-analysis had to evaluate storybook-focused reading (not a general literacy activity) and the storybooks had to be read by, or to, the children at school/lab or at home immediately following the intervention period. Additionally, eligible studies had a control group of some kind in the case of experiments (group comparison utilizing preexisting groups or within-subject design or randomized control trial design) and in the case of intervention, they had clear pre- and post-testing procedures.

Our systematic review only included children of certain characteristics. Specifically, participants were deemed eligible to be included in the analysis if they were typically developing children, of any gender and any social, cultural or geographical background, between the age of 2 and 10 years old. Research examining atypical developing individuals with developmental/acquired disorders, hearing impairments/deaf, bi-modal bilinguals were excluded. We made this decision given our knowledge of the field and the heterogeneity found in research investigating empathy development through storybooks for children with developmental disabilities. Of interest to this review were empirical primary studies that could be quantitatively systematized, whilst theoretical, conceptual, and review articles were first screened for additional references and then excluded. Studies with a unit of analysis or outcome variable in the form of one or several empathy-related skills were deemed eligible and included for further screening. Finally, the review focussed on studies written in English and included articles published in a peer-reviewed journal, unpublished dissertations, and pre-prints.

#### Literature search

The databases that we searched included ERIC/EBSCO, Academic Search Premier, APA PsycINFO, Scopus, Web of Science, MedLine Ovid, ProQuest, and PubMed. The systematic literature review followed detailed and pre-defined search terms (keywords determined through back literature search, expert discussions and search of relevant literature), inclusion and exclusion criteria developed using the PICO (population, intervention, comparison, and outcome) method to select relevant studies. We searched studies published between 1977 and 2022 in psychology and education databases. The latest search was conducted on 31 May 2022 to ensure that all the eligible studies were included in the review.

The search strategy was further scrutinized with a University Librarian at the (reference withheld) and is presented in Appendix A. Further manual examination of reference sections of the eligible studies was employed to seek additional studies. This review also aimed to include the so-called gray literature in the form of dissertations, unpublished studies, and conference proceedings. Where studies could not be obtained either in printed form or on-line by the reviewers, the authors were contacted via e-mail and asked to provide the papers for the purpose of the systematic review.

The results yielded by both automatic and manual searches were downloaded and formatted to enable further analyses with the use of Rayyan software (available at https://rayyan.ai/). The identified studies



Figure 2. PRISMA chart.

were uploaded onto Rayyan and screened for duplicates and irrelevant papers. The summary of this process along with the results yielded is presented in the PRISMA (Page et al., 2021) diagram in Figure 2.

Duplicates and studies that did not meet the inclusion criteria were removed. Full-text screening included 46 studies which were further assessed by three reviewers: each reviewer read the studies independently to determine their eligibility for inclusion. Our interrater reliability equaled  $\kappa = 0.86$ , which is considered very good. Two papers were not accessible online and therefore authors were contacted via e-mail and asked to provide full-text papers for the purpose of the systematic review. One author did not respond and thus their study was excluded as it was not available (a dissertation by Brockmeyer, 2009). The second author, Knotek (1996) replied but did not supply the specified paper.

The final 20 papers were included in further analyses. Studies reporting multiple empathy outcome measures were handled with robust variance estimation. The overall effect sizes are reported in Cohen's d and synthesized with a forest plot (with 95% confidence intervals).

#### **Quality assessment**

Once the set of eligible studies was finalized, the quality of the evidence was assessed by means of two separate quality assurance schemes. Gough's (2007) Weight of Evidence (WoE) criteria were employed. These criteria comprise four main questions (see Appendix B) and can be judged as high, medium, or low. WoE A concerns overall coherence and integrity of the individual study; WoE B refers to the appropriateness of the given form of evidence to answer a question; WoE C concerns the relevance of the evidence available to answer a question; WoE D refers to the overall

judgment of the evidence based on the previous question. Two reviewers independently assessed the quality of each study (see Table S1). In case of discrepancies between reviewers, the lowest assessment was accepted, as recommended by Murphy and Unthiah (2015).

Apart from Gough's WoE, we also employed an additional assessment protocol designed specifically to evaluate the Strength of Evidence based on Strength of Evidence framework by Murphy and Unthiah (2015) with evidence of high quality (where findings are highly secure and makes a substantial contribution to the existing evidence), medium (where findings are moderately secure and makes a contribution to the existing evidence), and low quality (where findings are insecure and add little to the existing evidence). The study features to which these criteria were applied are shown in Appendix B and the actual assessment of the set of the included studies is shown in Table S2.

The weight of the evidence measures were used to evaluate the studies retrieved through our systematic search. Only studies meeting robust evidence criteria were included in the review.

### **Coding procedure**

The data extraction was conducted by the first author. A coding procedure that specified the features of storybooks related to children's empathy-related skills was developed, with the following coding parameters: authors, publication year, status, and format (i.e., peer-reviewed, unpublished), sample size (as well as attrition and exclusion rates) and participants' age, gender, ethnicity, language, measures to assess empathy, allocation to condition (i.e., random, targeted), study design (i.e., within-or between-subject), points of assessment, intervention period (i.e., length of sessions), study location (i.e., school, home, or lab), study type (i.e., individual, pair, or group work), adult guidance (whether active, passive or not present), number of books included in the study, their format (i.e., paper versus digital), their content (i.e., different types of stories), outcome measures (empathy-related skills).

# Sample

21 studies (1 paper with two independent groups of participants - Biskin & Hoskisson, 1977) resulting in the total sample size of N = 2,293 children (N = 1,986 after exclusion and attrition rates were taken into consideration) were included in the review. Due to the inconsistent format of reporting sample age in some studies (i.e., ranges, use of school grades), it was not possible to calculate the exact mean age of participants. The estimated age mean calculated on the remaining studies which provided this information is M = 74.02 months (6 years 2 months), SD = 24.79. In total, 999 girls and 987 boys were included in the review (with two studies that did not report the gender distribution (Riquelme-Mella & García-Celay, 2016; Wulandini et al., 2017). 34 effect sizes for the between group (empathy emphasizing storybooks versus no empathy focused storybooks) difference in empathy-related skills were extracted. The extracted data comes from 10 countries; six from the United States, and one from each country: Jordan, Israel, China, Indonesia, Germany, Canada, Uganda, the United Kingdom, and South Korea. The included studies were published between 1977 and 2022. Where effect sizes were not reported, these were calculated based on the reported data (means, SDs, t-, F-values, and r) or by contacting the authors (e.g. Gil et al., 2014) to provide this information. Three studies (Batanova et al., 2016; Riquelme-Mella & García-Celay, 2016; Symons et al., 2005) were excluded from the analysis due to the lack of relevant data and no response from the authors. Sample characteristics are reported in Table 1.

Some studies comprised more than one experiment, and where this was the case and all experiments were of interest to the systematic review, the experiments were extracted and treated as separate studies. For example, Biskin and Hoskisson (1977) examined the effects of structured discussions of moral dilemmas in children's literature on moral reasoning. Both experiments had similar designs and stimuli, but they differed in the number of subjects tested and the frequency of the treatment. As both experiments were deemed eligible, they were recorded as two separate studies for the purpose of the analysis. Similarly, some studies included multiple outcome measures (i.e., Tompkins, 2015 investigated four separate empathy-

						Gender	
Study	Type of publication	Study design	Study location	Sample size	Age (in years and months)	distribution (females/males)	SES (using the terminology of the included studies)
Adrian et al.	journal	within	UK	34	4 y 2 m	15/19	working-class & middle-
Aram et al. (2017)	journal	between	Israel	61	5 y 10 m	30/31	low
Aram et al. (2017)	journal	within	Israel	61	5 y 10 m	30/31	low
Betawi (2022)	journal	between	Jordan	200	4–5 years old	95/105	private preschool
Bhavnagri and Samuels (1996)	journal	between	US	44	4 y 5 m	26/18	high
Biskin and Hoskisson (1977)	journal	between	US	20	10 y 10 m	8/12	not reported
Biskin and Hoskisson (1977)	journal	between	US	34	10 y 4 m	18/16	not reported
Du and Hao (2018)	journal	between	China	322	5 y 5 m	35/33,34/31, 35/33,33/30, 28/ 30	wage-earner family
Gil et al. (2014)	symposium paper	between	South Korea	24	б у	12/12	not reported
Goodman and Dent (2019)	book chapter	between	Uganda	123	4 y 10 m	70/53	low
Kruse et al. (2020)	journal	between	US	38	4 y 8 m	20/18	middle-class
Kumschick et al. (2014)	journal	between	Germany	208	7 y 11 m	127/81	not reported
Pelletier and Beatty (2015)	journal	between	Canada	186	4–5 years old	101/85	not reported
Russell and Cain (2022)	journal	between	UK	179	5 y 6 m	75/104	low
Salay (2018)	dissertation	between	US	108	б у	47/61	
Schapira and Aram (2020)	journal	within	Israel	50	4 y 8 m	24/26	mothers' education
Tompkins (2015)	journal	between	US	73	4 y 5 m	33/40	low
Wang et al. (2015)	journal	between	US	168	third and fourth graders	72/96	not reported
Wulandini et al. (2017)	book chapter	between	Indonesia	108	10 y 1 m	not reported	not reported

#### Table 1. Sample characteristics.

related outcome measures of interest to this review: false belief understanding, emotion understanding, social skills, and problem behaviors), and thus were included as separate experiments. Table 2 summarizes study characteristics including information regarding study design, measures used by researchers to assess empathy, intervention period (length and frequency of sessions where stated), intervention location, setting (group, dyad, or individualized reading), adult's role (whether active, passive or adult not present), number of books used in the intervention, their format, as well as outcome measures to assess empathy-related skills in children.

#### **Statistical analyses**

Given that Cohen's d is a widely recognized and extensively employed standardized measure of effect size (Kelley & Rausch, 2006), we converted all effect sizes into Cohen's d. Specifically, in studies with larger sample sizes (N > 20 per group), the bias inherent in Cohen's d diminishes,

		Intervention period			Adult's			
Author	Туре	(length of sessions)	Location	Setting	role	Number of books	Format	Outcome Measure
Adrian et al. (2005)	within	1 session	lab	dyad	adult active	4 stories	paper	false belief score
Aram et al. (2017)	between	2 sessions 1 week apart	home	dyad	adult active	1 book	paper	social understanding and prosocial behavior
Aram et al. (2017)	within	2 sessions 1 week apart	home	dyad	adult active	1 book	paper	social-emotional utterances
Betawi (2022)	between	4 weeks	school	group	adult active	4 stories	digital	moral integrity scale
Bhavnagri and Samuels (1996)	between	1 academic year	lab	individualised	adult active	15 books	paper	relationship enhancement
Biskin and Hoskisson (1977)	between	7 weeks x ~45 minutes	school	group	adult active	7 stories	paper	moral maturity scores (7 weeks)
Biskin and Hoskisson (1977)	between	18 weeks x ~45 minutes	school	group	adult active	18 stories	paper	moral maturity scores (18 weeks)
Du and Hao (2018)	between	2 sessions	school	individualised	no	5 stories	digital	donating behavior
Du and Hao (2018)	between	2 sessions	school	individualised	no	6 stories	digital	donating intentions
Gil et al. (2014)	between	1 session 30 minutes	school	individualised	adult active	1 story	digital	perspective-taking scores
Goodman and Dent (2019)	between	6 months (1 hour twice per week)	school	group	adult active	3–5 stories	paper	theory of mind scores
Kruse et al. (2020)		1 session	lab/school	individualised	adult active	3 stories	digital	sticker sharing
Kumschick et al. (2014)	between	8 weeks (2× 45-min lesson a week)	child centre	group	adult passive	1 book	paper	emotional vocabulary & explicit emotional knowledge
Pelletier and Beatty (2015)	between	1 session ~40 minutes	school	individualised	adult active	2 Aesop's fables	paper	theory of mind and fables understanding
Russell and Cain (2022)	between	2 sessions	school	individualised	adult active	Four story versions	paper	number of stickers shared pre- and post-story for each story condition
Salay (2018)	between & within	4 weeks (15 sessions)	school	group	adult active	5 picture books	paper	empathy scores & affective scores
Schapira and Aram (2020)	within	2 weeks (3 sessions)	home/ school	dyad	adult active	1 book	digital	empathy, prosocial attitude, & social coherence
Tompkins (2015)	between & within	5 weeks	child centre	individualised	adult active	15 commercially- available storybooks	paper	false belief understanding, emotion understanding, SSIS social skills, & SSIS problem behaviors
Wang et al. (2015)	between	5 sessions? (1 per week)	school	group	adult active	not specified	paper	prosocial behavior & social emotional assets
Wulandini et al. (2017)	between	5 consecutive days/ 20 minutes per day	school	individualised	no	(Literary fiction group) 3 books	paper	first-order, second-order, & total ToM test scores

#### Table 2. Study characteristics.

rendering its inherent simplicity a more compelling advantage. Following Cohen's scale (Cohen, 1988), an effect size of .2 was identified as a small effect, .5 was interpreted as a moderate effect, and .8 was referred to as a large effect. For the meta-analytic purposes, the effect sizes, reported as Cohen's d, were extracted from the studies and where these were not available, they were calculated based on the other values provided (means, standard deviations, correlations, Pearson r, F,- and t-values).

Our investigation into the relationship between storybook reading and empathy in children utilizes several statistical techniques. Subgroup analysis involves dividing studies based on pre-defined characteristics, such as storybook format (paper vs. digital) or reading location (school vs. lab). This allows us to examine if the overall effect size varies within these subgroups. Moderator analysis, a more formal approach, statistically assesses whether specific study characteristics (moderators) explain the observed variation in effect sizes across studies. Finally, meta-regression builds upon moderator analysis by using regression techniques to quantify the relationships between moderators and effect sizes, providing a more nuanced understanding of how different factors influence the effectiveness of storybook reading in promoting empathy. These combined approaches allow us to not only identify an overall effect but also explore how different aspects of storybook reading experiences might influence the development of children's empathy skills.

The analysis employed *robumeta* R package (Fisher et al., 2017) in R software as the majority of included studies comprised more than one effect size and thus a robust variance estimation (RVE) was the most suitable method. RVE enables the inclusion of effect sizes which are statistically dependent in a single meta-analysis. Additionally, RVE allows for information retainment without the need to investigate the underlying covariances of the effect sizes inputted for the analysis (Tanner-Smith et al., 2016). Addressing a critical limitation of both random and fixed effects models, Robust Variance Estimation (RVE) enables reliable meta-analytic inferences irrespective of non-independence among effect sizes. We used a correlated effects model with small-sample corrections and p was set to 0.8 (default for robumeta) as an estimation of correlation of within-sample effect sizes. For heterogeneity,  $I^2$  and tau<sup>2</sup> were run and are reported in the Results section.

The forest plot was screened visually for any potential outliers and additional heterogeneity measures were employed and are reported in the Results section.

#### Publication bias and robustness of results

The Egger Sandwich test for publication bias did not show a statistically significant effect (*intercept* = 0.46, p = .876, 95% CI [-5.41, -6.34,]). The analysis with different estimation correlation values (i.e., p = 0, 0.2, 0.4, 0.6, 0.8, and 1) for the dependent effect sizes revealed no significant differences in models. The results were identical up to three decimal places in all given calculations. All effect size estimates measured with similar precision should fall within the equivalence bound of d = -1.25 and d = 6.53.

#### Results

#### RQ1: the impact of storybook reading on children's empathy-related skills

Our first research question focussed on the impact of storybooks on children's empathy-related skills. The results of the meta-analysis of 17 studies (min = 1, mean = 3, median = 2, max = 16) suggest that there is a statistically significant effect of storybook reading on children's empathy skills (d = 0.72, SE = 0.28, p < .020, 95% CI [0.12, 1.32]). However, the between-study heterogeneity was very high ( $I^2 = 96.52\%$ ,  $tau^2 = 1.41$ ). The Forest plot in Figure 3 shows the impact of storybooks on children's empathy development. Similar to traditional forest plots, RVE forest plots provide point estimates of individual studies in a form of boxes whose dimensions are proportional to the weight assigned to each effect size. These are also shown in columns on the right along with the study design type. It should be noted that the weight is not proportional to the effect size variance of 95% CI in this case as in RVE the combined

	Forest Plot			
Studies		Effect Size	Weight	Study Type
Adrian et al. (2005)	-	0.563	0.738	within
Aram et al (2017)		0.000	0.100	
social understanding and prosocial behavior social-emotional utterances	•	0.637 -1.248	0.369 0.369	between within
Betawi (2022) moral integrity scale	-	2.891	0.749	between
Bhavnagri & Samuels (1996) relationship enhancement	-	1.090	0.715	between
Biskin & Hoskisson (1977) moral maturity scores (7 weeks) moral maturity scores (18 weeks)	- <u>+</u> -	0.686 1.056	0.341 0.341	between between
Du & Hao (2018) donating behaviour donating intentions	-	0.655 0.717	0.377 0.377	between between
Gil et al. (2014) perspective-taking scores	-#-	0.233	0.684	between
Goodman et al. (2019) theory of mind scores		0.169	0.754	between
Kruse, Faller, & Read (2021) sticker sharing (personalised stories) sticker sharing (nonpersonalised stories) sticker sharing (nonpersonalised stories) sticker sharing (nonpersonalised stories)		0.270 0.322 -0.070 -0.205	0.172 0.172 0.172 0.172	between between within within
Kumschick et al. (2014) emotional vocabulary explicit emotional knowledge identification of masked feelings	* +	3.317 3.049 6.528	0.244 0.244 0.244	between between between
Pelletier et al. (2015) theory of mind and fables understanding	•	0.883	0.759	between
Russell & Cain (2022) number of stickers shared pre- and post-story for each story condition		0.218	0.757	between
Salay (2018) empathy scores affective scores empathy scores affective scores	+ + +	-0.094 -0.344 0.063 0.892	0.188 0.188 0.188 0.188	between between within within
Schapira & Aram (2020)				
empathy prosocial attitude social coherence	-#- -#- -#-	0.676 0.676 0.629	0.242 0.242 0.242	within within within
Tompkins (2015) false belief understanding (experimenter-led book interactions) emotion understanding (experimenter-led book interactions) SSIS social skills (experimenter-led book interactions) SSIS problem behavious (experimenter-led book interactions) false belief understanding (storybook) SSIS problem behavious (exproybook) SSIS problem behavious (exproybook) false belief understanding (experimenter-led book interactions) false belief understanding (experimenter-led book interactions) SSIS problem behavious (storybook) SSIS problem behavious (experimenter-led book interactions) SSIS problem behavious (experimenter-led book interactions) false belief understanding (experimenter-led book interactions) false belief understanding (storybook) SSIS problem behavious (storybook) SSIS problem behavious (storybook)		0.591 0.458 0.511 -0.720 0.096 0.158 0.377 -0.493 -0.929 -0.867 -0.502 0.078 -0.188 -0.465 -0.284 0.194	0.045 0.045 0.045 0.045 0.045 0.045 0.045 0.045 0.045 0.045 0.045 0.045 0.045 0.045 0.045	between between between between between between between within within within within within within within
Wang et al. (2015) prosocial behavior social emotional assets		0.668 0.273	0.379 0.379	between between
Walandini et al. (2018) first-order ToM test scores (literary fiction) second-order ToM test scores (literary fiction) total TOM test scores (literary fiction) field TOM test scores (literary fiction) second-order ToM test scores (popular fiction) total TOM test scores (popular fiction)		0.950 0.321 0.950 0.259 0.238 0.259	0.121 0.121 0.121 0.121 0.121 0.121 0.121	between between between between between
	-2 0 2 4 6 8 Effect Size			

Figure 3. Forest plot for the impact of storybooks on empathy development in children.

study weight needs to be divided evenly between all the effect sizes in a study. This type of forest plot provides information regarding effect sizes for each individual study but at the same time it also enables presenting the study-level features as a part of the analysis. The overall effect size is depicted in a shape of a diamond with width equivalent to the 95% CI for the estimated effect size.

To address the high between-study heterogeneity in the main meta-analysis and our interest in individual empathy skills, we categorized each study in relation to the four empathy-related skills: Theory of Mind (ToM), prosocial behavior, emotional vocabulary/understanding, and perspective-taking (see Table 3).

# Table 3. Outcome measure categorization.

Author	ТоМ	prosocial behavior	emotional vocabulary/understanding	perspective-taking
Adrian et al. (2005)	false belief score			
Aram et al. (2017)		social understanding and prosoc	ial behavior	
Aram et al. (2017)			social-emotional utterances	
Betawi (2022)		moral integrity scale		
Bhavnagri and Samuels (1996)		relationship enhancement		
Biskin and Hoskisson (1977)		moral maturity scores		
Du and Hao (2018)		donating intentions and behavior		
Gil et al. (2014)	perspective-taking scores			
Goodman and Dent (2019)	theory of mind scores			
Kruse et al. (2020)		sticker sharing		
Kumschick et al. (2014)			emotional vocabulary, & explicit emotional knowledge	
Pelletier and Beatty (2015)	theory of mind and fables understanding		-	
Russell and Cain (2022)		stickers sharing		
Salay (2018)		empathy scores		
Salay (2018)			affective scores	
Schapira and Aram (2020)		empathy, prosocial attitude, & so	ocial coherence	
Tompkins (2015)	false belief understanding			
Tompkins (2015)		CCIC as sight skills, and much laws ha	emotion understanding	
Vang et al. (2015)		sola skills and problem be	ional assots	
Wulandini et al. (2017)	first-order second-order and total ToM tost		וטוומו מספרוס	
wulanumi et al. (2017)	scores			

The meta-analysis was not plotted for the perspective taking outcome measure as it only included one eligible study (Gil et al., 2014). We therefore ran separate meta-analyses for three of the four target empathy skills.

# Theory of mind (TOM) meta-analysis

Five studies with a total of 524 children (219 girls and 197 boys except Wulandini et al. (2017) where gender distribution is not reported) were included in the comparison of theory of mind scores between experimental (storybook reading) and control groups in between-subject designs as well as between pre- and posttest in within-subject designs (Figure 4). Two measures to assess the between-study heterogeneity were employed:  $I^2$  and tau<sup>2</sup> heterogeneity test.  $I^2$  test resulted in substantial hetero-geneity (78.62%). There was not a statistically significant effect size observed for either between- or within-subject performance on ToM, d = .42, 95% *CI* [-0.07; .91], p = .074, indicating that storybook reading has no statistically significant impact on children's theory of mind skills.

# Emotional vocabulary/understanding meta-analysis

Four studies with a total of 450 children (237 girls and 213 boys) were entered in the comparison of emotional vocabulary/understanding scores between experimental (storybook reading) and control groups in between-subject designs as well as between pre- and posttest in within-subject designs (Figure 5).

 $I^2$  test resulted in considerable heterogeneity (98.82%). Similarly to the ToM analysis, there was no statistically significant effect size observed for either between- or within-subject performance on emotional vocabulary/understanding, d = .51, 95% CI [-2.51; 3.53], p = .629.

# Prosocial behaviour meta-analysis with RVE

Twelve studies with a total of 1,297 children (613 girls and 684 boys) were included in the comparison of prosocial behavior scores between experimental (storybook reading) and control groups in



Figure 4. Forest plot for theory of mind performance.



Figure 5. Forest plot for emotional vocabulary/ understanding performance.

between-subject designs as well as between pre- and posttest in within-subject designs (Figure 6). Biskin and Hoskisson (1977) were entered as two separate studies due to the different sets of participants in each experiment.

 $I^2$  test resulted in considerable heterogeneity (93.58%). A statistically significant effect size was observed for between- and within-subject performance on prosocial behavior, d = .69, 95% CI [0.11; 1.26], p = .023, thus indicating that storybook reading significantly impacts children's prosocial behavior. Table 4 summarizes the results of the separate meta-analyses. Overall, the results of the above presented meta-analyses should be treated with caution due to a small number of eligible studies and high heterogeneity between the studies included in the analysis.

#### **RQ2: type of storybooks**

Our second Research Question focused on the format of the storybooks included in the review, and whether this is a factor that explains the beneficial effect. The moderator analysis revealed that format, whether the book is paper or digital, does not significantly affect empathy-related skills development in children (b = -0.07, SE = 0.21, p = .789, 95% CI [-2.24, 2.10]). Due to the missing information, studies not reporting and authors not responding to our requests for further information on the specific design and content of tested storybooks, we were unable to efficiently address the impact of content and interaction between storybook content and format on children's empathy.

#### **RQ3: type of reading**

Our third research question focused on the relationship between type of reading and the four empathy-related skills. The meta-regression analysis with Correlated and Hierarchical Effects (CHE) model was conducted using the whole sample with the following predictors: format of storybooks, adult's role (active, passive, or no presence), as well as the setting and the location of reading. With significance level set to p = .05, none of the variables of interest were found to be significant predictors. Where degrees of freedom are smaller than 4, these results should be treated with caution. As detailed in Table S3, individualized setting is the only predictor that is significant when robust variance estimation is used. This means that the effect sizes can be predicted by the group setting covariate.

#### Forest Plot

Studies								Effect Size	Weight	Study Type
Aram et al. (2017) social understanding and prosocial behavior			_	<b>-</b>				0.637	1.162	between
Betawi (2022) moral integrity scale								2.891	1.220	between
Bhavnagri & Samuels (1996) relationship enhancement				-	_			1.090	1.133	between
Biskin & Hoskisson (1977) moral maturity scores (7 weeks) moral maturity scores (18 weeks)				•				0.686 1.056	0.526 0.526	between between
Du & Hao (2018) donating behaviour donating intentions				*				0.655 0.717	0.616 0.616	between between
Kruse, Faller, & Read (2021) sticker sharing (personalised stories) sticker sharing (nonpersonalised stories) sticker sharing (personalised stories) sticker sharing (nonpersonalised stories)		_						0.270 0.322 -0.070 -0.205	0.266 0.266 0.266 0.266	between between within within
Russell & Cain (2022) number of stickers shared pre- and post-story for each story condition			-	F				0.218	1.242	between
Salay (2018) empathy scores empathy scores			-	-				-0.094 0.063	0.613 0.613	between within
Schapira & Aram (2020) empathy prosocial attitude social coherence			-	•				0.676 0.676 0.629	0.386 0.386 0.386	within within within
Tompkins (2015) SSIS social skills (experimenter-led book interactions) SSIS problem behaviours (experimenter-led book interactions) SSIS social skills (storybook) SSIS problem behaviours (storybook) SSIS problem behaviours (experimenter-led book interactions) SSIS problem behaviours (experimenter-led book interactions) SSIS problem behaviours (storybook) SSIS problem behaviours (storybook)				+				0.511 -0.720 0.377 -0.493 -0.502 0.078 -0.284 0.194	0.145 0.145 0.145 0.145 0.145 0.145 0.145 0.145 0.145	between between between within within within within
Wang et al. (2015) prosocial behavior social emotional assets			-1	+				0.668 0.273	0.623 0.623	between between
	2	1		1	2	- 1				
	-2	-1	E	ffect Si	ze	3	4			

Figure 6. Forest plot for prosocial behavior performance.

Table 4. Subgroup a	nalysis summary.
---------------------	------------------

Outcome Measure	d	95% CI	р	l <sup>2</sup>
ToM	0.42	-0.07; .91	0.074	78.62
emotional vocabulary/understanding	0.51	-2.51; 3.53	0.629	98.82
prosocial behaviour*	0.69	0.11; 1.26	0.023	93.58

\*Significant at the p < .05.

In other words, when children read on their own or in a one-to-one reading session, there is more likely to be a gain in empathy skills as a result storybook reading, compared to when reading in a group.

Given that three studies (Betawi, 2022; Pelletier & Beatty, 2015; Wang et al., 2015) did not report exact means for age, only age ranges, the analysis of possible differential impact of storybook reading on the empathy skills of children at different age, could not be carried out reliably for the whole sample. We calculated an age average for Betawi (2022) and Pelletier and Beatty (2015) studies (i.e. where the authors report an age range of 4–5 years, a mean age of 4.5 was computed) and added to the analysis. Wang et al. (2015) was excluded due to the difficulties in estimating a reliable mean age based on the information provided by the authors. Based on these data, we checked the significance of age as a predictor. This analysis was not significant: [b = -0.179, SE = 0.027, df = 1.35, p = .060]. The meta-regression plotted with the exclusion of the three studies stated above did not change the significance of age as a predictor: [b = -0.151, SE = 0.098, df = 1.19, p = .339].

#### Discussion

This study investigated the impact of storybook reading on a range of children's empathy-related skills, across various book types and across various reading sessions. We also explored the potentially moderating effects of storybooks' format, adult's role (active, passive, or no presence), as well as the setting and the location of reading.

We found that overall, storybook reading has a significant effect on empathy skills, but this effect only holds for prosocial behavior. The overall impact appears medium (d = 0.72, p < .020) but can be considered large in education research (i.e. Cohen's d > 0.6). However, the between-study heterogeneity that we report is extremely high ( $I^2 = 96.52\%$ ) which might be due to various experimental setups between included studies, different ways of assessing empathy, and different definitions of empathy itself. The additional step of dividing the data into four empathy-related categories allows a more skillspecific insight into where the true effect size is found (prosocial behavior) but it also weakens the overall effect calculations. While the format of storybooks does not seem to affect the impact of the reading session on children's empathy (paper: b = -0.13, t(1) = -0.43, p = .731), the type of reading context does. Namely, individualized, as opposed to dyadic reading with parents or group reading with peers, was found to be most relevant for empathy skills (b = 0.53, t(5) = -0.074, p < .001). It needs to be noted that in the subgroup analyses, the results for Theory of Mind (TOM) and Emotional Vocabulary/ Understanding were insignificant, yet considerable heterogeneity was within these subgroups. Moreover, the number of studies included in each subgroup was fewer than five, which could have resulted in misleading conclusions (Guolo & Varin, 2017). Egger's test indicated no evidence of publication bias. We discuss these findings with attention to their implications for interdisciplinary research on storybook reading and empathy and corresponding policy-making recommendations.

Our findings highlight an interesting discrepancy – storybook reading significantly impacts prosocial behavior but not necessarily other facets of empathy. This selective influence might be attributed to the inherent nature of prosocial behavior. Stories often depict characters navigating social situations, overcoming challenges, and helping others. This focus on action and consequence may resonate more strongly with children, translating directly into increased willingness to help themselves. Meanwhile, aspects like emotional understanding, which involve a deeper grasp of complex feelings, might require a more nuanced approach beyond simple story exposure. Future research could explore if targeted discussions or activities alongside story time can bridge this gap, fostering a more comprehensive development of empathy in children.

#### Study implications

Taking an interdisciplinary viewpoint, engaging in storybook reading is likely to enhance children's empathy skills, given that the activity generally incorporates various attributes that have the potential to promote empathy. The role of these attributes in empathy is disciplinary divided. From a psychology perspective, Mar and Oatley (2008) theorize that it is the narrative, which is present in storybooks but also films and plays, that simulates experience and explains increased perspective-taking skills. From a literary perspective, the seminal work by Nikolajeva on children's reading of fiction (2009) posits that the fictional nature of characters and stories (e.g. personified animals,

superheroes or witches) supports empathy because it allows young readers to practice their awareness of what people in different situations might feel or experience. Yet another perspective comes from the self-distancing theory, which proposes that experiences, such as storybook reading, achieve their benefits through the mental distance they introduce between the "self" and the "other" (Orvell et al., 2019). The different causal mechanisms constitute possible explanatory models for our findings.

Namely, if it is the presence of fictional narrative that leads to higher empathy skills then arguably, the format of the book does not matter in empathy gains and storybook reading interventions could be delivered in both paper and digital book reading formats. However, given that we could not examine the impact of content on children's empathy development, the content of the storybooks remains to be investigated. Different types of content features (e.g. story plots and moral of the story, personal characteristics and depiction of diverse story characters) offer different possibilities for experiencing the "self" and the "other" in fictional narratives. The role of these elements in children's storybooks is important to determine the type of empathy being promoted by storybooks. The experience of reading about story characters who are different from the child, and therefore members of the out-group, is hypothesized to constitute the highest empathy potential of children's storybooks (Kucirkova, 2019) but there is minimal empirical work to support this theoretical proposition. Our findings re-introduce Burke et al. (2016) recommendation for the empathy field to attribute particular elements of storybooks to particular segments of empathy, in order to allow an evidence-based understanding of the association between the benefits of storybook reading and empathy.

The finding that enhanced empathy is seen in individualized reading contexts, but not group-based or parent-child dyadic reading, indicates the importance of context, both in how empathy is developed, but also, perhaps, in how it is assessed. Although the exact details of this condition varied across studies, a common theme was that a child would read and interact with a potentially unfamiliar adult who would be giving individualized prompts and encouragement. Such a socially positive and connecting context could optimize both receptivity to the storybook content, as well as assessment performance. While previous research has looked at the role of extra-textual talk in developing empathic ability (Ornaghi & Grazzani, 2013), it may also be important to consider who the talkers are, and what the social relationship is between them.

So that future empathy intervention studies can better delineate the impact of storybook reading and more strongly influence children's empathy skills, we suggest more targeted, longer and intense (increased frequency) storybook reading interventions. Another methodological recommendation concerns the choice of control groups: future interventions could mirror adult empathy reading studies with control groups of nonfiction, or alternative activities involving fictional narratives (e.g. films and role-play). Finally, the interventions examined in the present meta-analysis focused on children's immediate gains in empathy skills. While this provides a window into dispositional and situational effects that are related to empathy, it does not explain long-term effects that may extend beyond intervention periods.

#### **Study limitations**

Our meta-analysis was limited with respect to the diversity of participants and research teams, with a focus on studies published in English. Furthermore, we focused only on typically developing children, which limits the generalizability of our findings. Our analysis could be extended with a narrative review focused on clinical populations and the population of students who often require explicit instruction in social/emotional learning (e.g. children with autism). Future studies should build on the study of empathy effects in diverse reading settings, with children from various socio-economic, cultural and ethnic backgrounds and various abilities. Due to an insufficient number of studies, we could not examine effects of gender or age on the overall empathy impact – an issue to be taken up by later meta-analyses when more experimental data becomes available. To address this limitation in future research, it is recommended to conduct a power analysis apriori (Griffin, 2021) to determine the minimum number of studies required to investigate the moderating effects of gender, age, or other relevant factors.

The storybook format was not found to be a significant predictor in the meta regression analysis when robust variance estimation was used. In other words, in the studies included in this analysis, there was no difference between the paper and digital version of the books used in interventions. This connects to Kucirkova's (2019) proposition that as long as storybooks contain a fictional narrative, they can be in any format (digital or paper-based) and presented verbally or visually, to foster children's empathy skills. Future studies could usefully supplement this finding with an examination of the various features of print and digital books. Notably the role of interactivity and the possibility to directly move characters on the screen in digital books was theorized to be relevant for children's empathy (Zhao & Unsworth, 2016), but could not be examined in our meta-analysis due to the lack of available detail concerning the design of the storybooks.

Significant impacts of book-reading were observed only for prosocial behaviors – a domain in which there were at least twice as many studies as in any of the other three domains. Caution should be exercised when interpreting the strength of the evidence. Furthermore, the prosocial outcomes were measured in rather divergent ways (donating, sharing stickers, moral maturity or integrity scale, empathy scores, social understanding), raising the possibility that they may have not been tapping the same basic construct. We recommend that future research interested in children's storybook reading and empathy strategically targets prosocial skills as the most pertinent empathy-related area to be refined and expanded through empirical research.

Finally, although our meta-analysis provides some answers, many unanswered questions remain: for example, children's literature scholars are likely to be interested in the quality markers that characterize empathy-promoting storybooks. Our meta-analysis did not provide this information as it could only draw on experimental studies with their specific methods, many of which were not fully standardized.

#### Conclusion

A systematic examination of empathy-relevant content and format features of storybooks is necessary to understand the learning mechanisms behind "what works" for theory of mind, perspective-taking, and emotional literacy and intelligence, and for informing adults' choices of children's books in multiple formats. There is not enough evidence to determine the impact of the format of storybooks on empathy. Given the wider accessibility of digital books on family smartphones and home devices, e-reading empathy interventions could offer a cost-effective alternative to print books. An important avenue for future studies is thus to analyze the features and characteristics of storybooks that were used in intervention studies. Literature remains still unclear on whether empathy is promoted by all types of storybooks or only books with a specific content. Overall, storybook reading is mostly related to prosocial behavior and this finding adds a vital nuance to the guidance and criteria about empathyenabling features of children's storybooks.

#### **Disclosure statement**

No potential conflict of interest was reported by the author(s).

#### References

- Abramson, L., Uzefovsky, F., Toccaceli, V., & Knafo-Noam, A. (2020). The genetic and environmental origins of emotional and cognitive empathy: Review and meta-analyses of twin studies. *Neuroscience & Biobehavioral Reviews*, 114, 113–133. https://doi.org/10.1016/j.neubiorev.2020.03.023
- \*Adrian, J. E., Clemente, R. A., Villanueva, L., & Rieffe, C. (2005). Parent-child picture-book reading, mothers' mental state language and children's theory of mind. *Journal of Child Language*, 32(3), 673–686. https://doi.org/10.1017/ s0305000905006963
- \*Aram, D., Bergman Deitcher, D., Sabag Shoshan, T., & Ziv, M. (2017). Shared book reading interactions within families from low socioeconomic backgrounds and children's social understanding and prosocial behavior. *Journal of Cognitive Education & Psychology*, 16(2), 157–177. https://doi.org/10.1891/1945-8959.16.2.157

Baron-Cohen, S. (2011). Zero degrees of empathy: A new theory of human cruelty. Penguin UK.

- Bar-Tal, D., Raviv, A., & Goldberg, M. (1982). Helping behavior among preschool children: An observational study. Child Development, 53(2), 396–402. https://doi.org/10.2307/1128982
- Batanova, M., Bowers, E. P., Hilliard, L. J., Tirrell, J. M., Stacey, D. C., McClain, A., & Lerner, R. M. (2016). Examining cross-age peer conversations relevant to character: Can a digital story about bullying promote students' understanding of humility? *Research in Human Development*, 13(2), 111–125. https://doi.org/10.1080/15427609.2016. 1166014
- Berg, T. (1998). Linguistic structure and change: An explanation from language processing. Oxford University Press.
- Bernard, S., & Deleau, M. (2007). Conversational perspective-taking and false belief attribution: A longitudinal study. British Journal of Developmental Psychology, 25(3), 443–460. https://doi.org/10.1348/026151006X171451
- \*Betawi, I. A. (2022). Children for integrity: The influence of moral stories on pre-schoolers' integrity values. *Clinical Child Psychology and Psychiatry*, 28(2), 541–553. https://doi.org/10.1177/13591045221094694
- \*Bhavnagri, N., & Samuels, B. (1996). Children's literature and activities promoting social cognition of peer relationships in preschoolers. *Early Childhood Research Quarterly*, 11(3), 307–331. https://doi.org/10.1016/S0885-2006(96)90010-1 Bierhoff, H. W. (2002). *Prosocial behaviour*. Psychology Press.
- \*Biskin, D. S., & Hoskisson, K. (1977). An experimental test of the effects of structured discussions of moral dilemmas found in children's literature on moral reasoning. *The Elementary School Journal*, 77(5), 407–416. https://doi.org/10. 1086/461074
- Bjorkqvist, K., Osterman, K., & Kaukiainen, A. (2000). Social intelligence-empathy = aggression? Aggression and Violent Behavior, 5(2), 191–200. https://doi.org/10.1016/s1359-1789(98)00029-9
- Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition*, 14(4), 698–718. https://doi.org/10.1016/j.concog.2005.06. 004
- Bloom, P. (2017). Against empathy: The case for rational compassion. Random House.
- Brockmeyer, C. (2009). Re-reading the relationship between narrative and theory of mind: The effects of narrative training on the developing understanding of mind. Lehigh University ProQuest Dissertations Publishing.
- Brown, S., Harkins, L., & Beech, A. R. (2012). General and victim-specific empathy: Associations with actuarial risk, treatment outcome, and sexual recidivism. Sexual Abuse: Journal of Research and Treatment, 24(5), 411–430. https:// doi.org/10.1177/1079063211423944
- Bruneau, E. G., Cikara, M., & Saxe, R. (2015). Minding the gap: Narrative descriptions about mental states attenuate parochial empathy. PLOS ONE, 10(10), e0140838. https://doi.org/10.1371/journal.pone.0140838
- Bruneau, E. G., Cikara, M., & Saxe, R. (2017). Parochial empathy predicts reduced altruism and the endorsement of passive harm. Social Psychological and Personality Science, 8(8), 934–942. https://doi.org/10.1177/1948550617693064
- Burke, M., Kuzmičová, A., Mangen, A., & Schilhab, T. (2016). Empathy at the confluence of neuroscience and empirical literary studies. *Scientific Study of Literature*, 6(1), 6–41. https://doi.org/10.1075/ssol.6.1.03bur
- Cassidy, K. W., Ball, L. V., Rourke, M. T., Werner, R. S., Feeny, N., Chu, J. Y., Perkins, A., & Perkins, A. (1998). Theory of mind concepts in children's literature. *Applied Psycholinguistics*, 19(3), 463–470. https://doi.org/10.1017/ S0142716400010274
- Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and them: Intergroup failures of empathy. *Current Directions in Psychological Science*, 20(3), 149–153. https://doi.org/10.1177/0963721411408713
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates, Publishers.
- Cox, C. L., Uddin, L. Q., DiMartino, A., Castellanos, F. X., Milham, M. P., & Kelly, C. (2012). The balance between feeling and knowing: Affective and cognitive empathy are reflected in the brain's intrinsic functional dynamics. Social Cognitive and Affective Neuroscience, 7(6), 727–737. https://doi.org/10.1093/scan/nsr051
- Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review*, 8(2), 144–153. https://doi.org/10.1177/1754073914558466
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. Journal of Personality and Social Psychology, 44(1), 113–126. https://doi.org/10.1037/0022-3514.44.1.113
- De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, 14(10), 441-447. https://doi.org/10.1016/j.tics.2010.06.009
- Denham, S. A., Bassett, H. H., Thayer, S. K., Mincic, M. S., Sirotkin, Y. S., & Zinsser, K. (2012). Observing Preschoolers'Social-emotional behavior: Structure, foundations, and prediction of early school success. *The Journal* of Genetic Psychology: Research and Theory on Human Development, 173(3), 246–278.
- Dodell-Feder, D., & Tamir, D. I. (2018). Fiction reading has a small positive impact on social cognition: A meta-analysis. *Journal of Experimental Psychology*, 147(11), 1713. https://doi.org/10.1037/xge0000395
- \*Du, X., & Hao, J. (2018). Moral stories emphasizing actors' negative emotions toward their nonhelping behavior promote preschoolers' helping behavior. *Journal of Experimental Child Psychology*, *168*, 19–31. https://doi.org/10. 1016/j.jecp.2017.11.011

- Fikrat-Wevers, S., van Steensel, R., & Arends, L. (2021). Effects of family literacy programs on the emergent literacy skills of children from low-SES families: A meta-analysis. *Review of Educational Research*, 91(4), 577–613. https://doi.org/ 10.3102/0034654321998075
- Fisher, Z., Tipton, E., & Zhipeng, H. (2017). Robumeta (Version 2.0). [Computer Software]. https://cran.r-project.org/ package=robumeta
- Furenes, M. I., Kucirkova, N., & Bus, A. G. (2021). A comparison of children's reading on paper versus screen: A metaanalysis. *Review of Educational Research*, 91(4), 0034654321998074. https://doi.org/10.3102/0034654321998074
- Garner, P. W., & Estep, K. M. (2001). Emotional competence, emotion socialization, and young children's peer-related social competence. *Early Education and Development*, 12(1), 29–48. https://doi.org/10.1207/s15566935eed1201\_3
- \*Gil, K., Rhim, J., Ha, T., Doh, Y. Y., & Woo, W. (2014). AR petite theater: Augmented reality storybook for supporting children's empathy behavior. In 2014 IEEE International Symposium on Mixed and Augmented Reality - Media, Art (pp. 13–20). Social Science, Humanities and Design (ISMAR-MASH'D). https://doi.org/10.1109/ISMAR-AMH.2014. 6935433
- Gloor, J. L., & Puhl, R. M. (2016). Empathy and perspective-taking: Examination and comparison of strategies to reduce weight stigma. *Stigma and Health*, 1(4), 269. https://doi.org/10.1037/sah0000030
- \*Goodman, G., & Dent, V. F. (2019). Studying the effectiveness of the storytelling/story-acting (STSA) play intervention on Ugandan preschoolers' emergent Literacy, oral language, and theory of mind in two rural Ugandan community libraries. *Early Childhood Development*. https://doi.org/10.4018/978-1-5225-7507-8.ch059
- Gottschall, J. (2012). The storytelling animal: How stories make us human. Houghton Mifflin Harcourt.
- 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. https://doi.org/10.1080/02671520701296189
- Griffin, J. W. (2021). Calculating statistical power for meta-analysis using metapower. The Quantitative Methods for Psychology, 17(1), 24–39. https://doi.org/10.20982/tqmp.17.1.p024
- Guolo, A., & Varin, C. (2017). Random-effects meta-analysis: The number of studies matters. Statistical Methods in Medical Research, 26(3), 1500–1518. https://doi.org/10.1177/0962280215583568
- Healey, M. L., & Grossman, M. (2018). Cognitive and affective perspective-taking: Evidence for shared and dissociable anatomical substrates. *Frontiers in Neurology*, 9, 491. https://doi.org/10.3389/fneur.2018.00491
- Herrera, F., Bailenson, J., Weisz, E., Ogle, E., & Zaki, J. (2018). Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PLOS ONE*, *13*(10), e0204494. https://doi.org/10.1371/journal.pone. 0204494
- Hoffman, M. L. (2000). Empathy and moral development: Implications for caring and justice. Cambridge University Press.
- Izard, C., Fine, S., Schultz, D., Mostow, A., Ackerman, B., & Youngstrom, E. (2001). Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Erratum In: Psychological Science*, 12(1), 18–23. https:// doi.org/10.1111/1467-9280.00304
- Keen, S. (2007). Empathy and the novel. Oxford University Press.
- Kelley, K., & Rausch, J. R. (2006). Sample size planning for the standardized meandifference: Accuracy in parameter estimation via narrow confidence intervals. *Psychological Methods*, 11(4), 363–385. https://doi.org/10.1037/1082-989X.11.4.363
- Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. Science, 342(6156), 377–380. https:// doi.org/10.1126/science.1239918
- Knotek, S. E. (1996). Emergent literacy and theory of mind: Children's use of text as situation and representation to mediate social interactions [PhD dissertation, University of California].
- Koopman, E. M. (2018). Does originality evoke understanding? The relation between literary reading and empathy. *Review of General Psychology*, 22(2), 169–177. https://doi.org/10.1037/gpr0000107
- \*Kruse, E., Faller, I., & Read, K. (2020). Can reading personalized storybooks to children increase their prosocial behavior? *Early Childhood Education Journal*, 49(2), 273–282. https://doi.org/10.1007/s10643-020-01069-x
- Kucirkova, N. (2019). How could children's storybooks promote empathy? A conceptual framework based on developmental psychology and literary theory. *Frontiers in Psychology*, 10, 121. https://doi.org/10.3389/fpsyg.2019.00121
- Kucirkova, N., Snow, C. E., Grøver, V., & McBride, C. (Eds.). (2017). The Routledge international handbook of early literacy education: A contemporary guide to literacy teaching and interventions in a global context. Taylor & Francis.
- Kümmerling-Meibauer, B. (2014). What goes on in strangers' minds? How reading children's books affects emotional development. Narrative Works: Issues, Investigations, and Interventions, 4(2). https://doi.org/10.7202/1062100ar
- \*Kumschick, I. R., Beck, L., Eid, M., Witte, G., Klann-Delius, G., Heuser, I., Steinlein, R., & Menninghaus, W. (2014). Reading and Feeling: The effects of a literature-based intervention designed to increase emotional competence in second and third graders. *Frontiers in Psychology*, 5, 1448. https://doi.org/10.3389/fpsyg.2014.01448
- Lee, K., Talwar, V., McCarthy, A., Ross, I., Evans, A., & Arruda, C. (2014). Can classic moral stories promote honesty in children? *Psychological Science*, 25(8), 1630–1636. https://doi.org/10.1177/0956797614536401
- Luke, N., & Banerjee, R. (2013). Differentiated associations between childhood maltreatment experiences and social understanding: A meta-analysis and systematic review. *Developmental Review*, 33(1), 1–28. https://doi.org/10.1016/j. dr.2012.10.001

- Mar, R. A. (2011). The neural bases of social cognition and story comprehension. Annual Review of Psychology, 62(1), 103–134. https://doi.org/10.1146/annurev-psych-120709-145406
- Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. Perspectives on Psychological Science, 3(3), 173–192. https://doi.org/10.1111/j.1745-6924.2008.00073.x
- Mar, R. A., Oatley, K., Hirsh, J., Dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Jornal of Research* in Personality, 40, 694–712.
- Mar, R. A., Tackett, J. L., & Moore, C. (2010). Exposure to media and theory-of-mind development in preschoolers. Cognitive Development, 25(1), 69–78. https://doi.org/10.1016/j.cogdev.2009.11.002
- Mason, J. M. (1990). Shared Book Reading in an Early Start Program for at-Risk Children. Center for the study of reading technical Report; no. 504.
- McDonald, N. M., & Messinger, D. S. (2011). The development of empathy: How, when, and why. In A. Acerbi, J. A. Lombo, & J. J. Sanguineti (Eds.), *Free will, emotions, and moral actions: Philosophy and neuroscience in dialogue* (pp. 333–359). IF-Press.
- Mitra, S., & Crawley, E. (2014). Effectiveness of self-organised learning by children: Gateshead experiments. *Journal of Education & Human Development*, 3(3), 79–88. https://doi.org/10.15640/jehd.v3n3a6
- Mumper, M. L., & Gerrig, R. J. (2017). Leisure reading and social cognition: A meta-analysis. *Psychology of Aesthetics, Creativity, and the Arts*, 11(1), 109–120. https://doi.org/10.1037/aca0000089
- Murphy, V., & Unthiah, A. (2015). A systematic review of intervention research examining English language and literacy development in children with English as an additional language (EAL). www.naldic.org.uk/Resources/NALDIC/ ResearchandInformation/Documents/eal-systematic-review-prof-v-murphy.pdf
- Neuman, S. B., Samudra, P., & Wong, K. M. (2021). Two may be better than one: Promoting incidental word learning through multiple media. *Journal of Applied Developmental Psychology*, 73, 101252. https://doi.org/10.1016/j.appdev. 2021.101252
- Nikolajeva, M. (2009). Power, voice and subjectivity in literature for young readers. Routledge.
- Nikolajeva, M. (2012). Guilt, empathy and the ethical potential of children's literature. *Barnboken*, 35(1), 18081. https://doi.org/10.3402/clr.v35i0.18081
- Noble, C., Sala, G., Peter, M., Lingwood, J., Rowland, C., Gobet, F., & Pine, J. (2019). The impact of shared book reading on children's language skills: A meta-analysis. *Educational Research Review*, 28, 100290. https://doi.org/10.1016/j. edurev.2019.100290
- Nyhout, A., & O'Neill, D. K. (2013). Mothers' complex talk when sharing books with their toddlers: Book genre matters. *First Language*, 33(2), 115–131. https://doi.org/10.1177/0142723713479438
- Oatley, K. (2016). Fiction: Simulation of social worlds. *Trends in Cognitive Sciences*, 20(8), 618–628. https://doi.org/10. 1016/j.tics.2016.06.002
- Ornaghi, V., Brockmeier, J., & Grazzani, I. (2014). Enhancing social cognition by training children in emotion understanding: A primary school study. *Journal of Experimental Child Psychology*, 119, 26–39. https://doi.org/10.1016/j. jecp.2013.10.005
- Ornaghi, V., & Grazzani, I. (2013). The relationship between emotional-state language and emotion understanding: A study with school-age children. *Cognition & Emotion*, 27(2), 356–366. https://doi.org/10.1080/02699931.2012.711745
- Orvell, A., Kross, E., & Gelman, S. A. (2019). Lessons learned: Young children's use of generic-you to make meaning from negative experiences. *Journal of Experimental Psychology: General*, 148(1), 184–191. https://doi.org/10.1037/ xge0000445
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., & Mulrow, C. D., (2021). The PRISMA, 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, 10(1), 10–89. https://doi.org/ 10.31222/osf.io/v7gm2
- Panero, M. E., Weisberg, D. S., Black, J., Goldstein, T. R., Barnes, J. L., Brownell, H., & Winner, E. (2016). Does reading a single passage of literary fiction really improve theory of mind? An attempt at replication. *Journal of Personality and Social Psychology*, 111(5), e46–e54. https://doi.org/10.1037/pspa0000064
- Pearson, J., Owen, T., Thimbleby, H. W., & Buchanan, G. (2012). Co-reading: Investigating collaborative group reading. In K. B. Boughida, B. Howard, & M. L. Nelson (Eds.), JCDL '12 Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries. 12th ACM/IEEE-CS joint conference on Digital Libraries, Washington, DC, USA, 10-14 Jun 2012. https://doi.org/10.1145/2232817.2232876
- \*Pelletier, J., & Beatty, R. (2015). Children's understanding of Aesop's fables: Relations to reading comprehension and theory of mind. *Frontiers in Psychology*, 6, 1448. https://doi.org/10.3389/fpsyg.2015.01448
- Rehfuess, E. A., Booth, A., Brereton, L., Burns, J., Gerhardus, A., Mozygemba, K., Oortwijn, W., Pfadenhauer, L. M., Tummers, M., van der Wilt, G.-J., & Rohwer, A. (2018). Towards a taxonomy of logic models in systematic reviews and health technology assessments: A priori, staged, and iterative approaches. *Research Synthesis Methods*, 9(1), 13–24. https://doi.org/10.1002/jrsm.1254
- Riquelme-Mella, E., & García-Celay, I. M. (2016). Long-term effects of a mediated reading programme on the development of emotional competencies. *Culture and Education*, 28(3), 435–467. https://doi.org/10.1080/11356405. 2016.1196900

- Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mother's mental state language and theory-of-mind understanding. *Child Development*, 73(3), 734–751. https://doi.org/10.1111/1467-8624.00435
- \*Russell, S. J., & Cain, K. (2022). The animals in moral tales: Does character realism influence children's prosocial response to stories? *Journal of Experimental Child Psychology*, 219, 105392. https://doi.org/10.1016/j.jecp.2022.105392
- \*Salay, D. M. (2018). Walk in their shoes: How picture books and critical literacy instruction can foster empathy in first grade Students.
- Saxe, R. (2006). Uniquely human social cognition. Current Opinion in Neurobiology, 16(2), 235–239. https://doi.org/10. 1016/j.conb.2006.03.001
- \*Schapira, R., & Aram, D. (2020). Shared book reading at home and preschoolers' socio-emotional Competence. *Early Education and Development*, 31(6), 819–837. https://doi.org/10.1080/10409289.2019.1692624
- Schapira, R., Bergman Deitcher, D., & Aram, D. (2021). Mothers' discourse during shared reading of books relating to 'positive' and 'negative' emotions in different genres. *Journal of Research in Reading*, 44(4), 897–909. https://doi.org/ 10.1111/1467-9817.12370
- Sénéchal, M. (2017). Shared book reading: An informal literacy activity par excellence. In N. Kucirkova, C. E. Snow, V. Grøver, & C. McBride (Eds.), *The Routledge international handbook of early literacy education* (pp. 273–283). Routledge.
- Silvera, D. H., Martinussen, M., & Dahl, T. I. (2001). The Tromso social intelligence scale, a self-report measure of social intelligence. *Scandinavian Journal of Psychology*, 42(4), 313–319. https://doi.org/10.1111/1467-9450.00242
- Spinrad, T. L., & Gal, D. E. (2018). Fostering prosocial behavior & empathy in young children. Current Opinion in Psychology, 20, 40–44. https://doi.org/10.1016/j.copsyc.2017.08.004
- Sulzby, E. (1985). Children's emergent reading of favorite storybooks: A developmental study. *Reading Research Quarterly*, 458-481. https://doi.org/10.1598/RRQ.20.4.4 20 (4)
- Symons, D. K., Peterson, C. C., Slaughter, V., Roche, J., & Doyle, E. (2005). Theory of mind and mental state discourse during book reading and story-telling tasks. *British Journal of Developmental Psychology*, 23(1), 81–102. https://doi. org/10.1348/026151004X21080
- Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks: A meta-analysis. *Review of Educational Research*, 85(4), 698–739. https://doi.org/ 10.3102/0034654314566989
- Tamir, D. I., Bricker, A. B., Dodell-Feder, D., & Mitchell, J. P. (2016). Reading fiction and reading minds: The role of simulation in the default network. Social Cognitive and Affective Neuroscience, 11(2), 215–224. https://doi.org/10. 1093/scan/nsv114
- Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. Annual Review of Psychology, 58 (1), 345. https://doi.org/10.1146/annurev.psych.56.091103.070145
- Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and life-Course Criminology*, 2(1), 85–112. https://doi. org/10.1007/s40865-016-0026-5
- \*Tompkins, V. (2015). Improving low-income preschoolers' theory of mind: A training study. *Cognitive Development*, 36, 1–19. https://doi.org/10.1016/j.cogdev.2015.07.001
- Vaish, A., Carpenter, M., & Tomasello, M. (2016). The early emergence of guilt-motivated prosocial behavior. *Child Development*, 87(6), 1772–1782. https://doi.org/10.1111/cdev.12628
- Vega, N., Stanfield, J., & Mitra, S. (2020). Investigating the impact of computer supported collaborative learning (CSCL) to help improve reading comprehension in low performing urban elementary schools. *Education & Information Technologies*, 25(3), 1571–1584. https://doi.org/10.1007/s10639-019-10023-3
- \*Wang, C., Couch, L. K., Rodriguez, G., & Lee, C. M. (2015). The bullying literature project: Using Children's literature to promote prosocial behavior and social-emotional outcomes among elementary school Students. *Contemporary School Psychology*, 19(4), 320–329. https://doi.org/10.1007/s40688-015-0064-8
- Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development*, 69(3), 848–872. https://doi.org/10.1111/j.1467-8624.1998.tb06247.x
- \*Wulandini, Kuntoro, I. A., & Handayani, E. (2017). The effect of literary fiction on school-aged children's theory of mind (ToM). Diversity in Unity: Perspectives from Psychology and Behavioral Sciences, 159–166. https://doi.org/10. 1201/9781315225302-20
- Yang, D., Xia, C., Collins, P., & Warschauer, M. (2022). The role of bilingual discussion prompts in shared E-book reading. Computers & Education, 190, 104622. https://doi.org/10.1016/j.compedu.2022.104622
- Yu, C. L., & Chou, T. L. (2018). A dual route model of empathy: A neurobiological prospective. Frontiers in Psychology, 9. https://doi.org/10.3389/fpsyg.2018.02212
- Zhao, S., & Unsworth, L. (2016). Touch design and narrative interpretation: A social semiotic approach to picture book apps. In N. Kucirkova & G. Fallon (Eds.), Apps, technology and younger learners (pp. 109–121). Routledge.

# Appendix A - Search terms used in the Literature Search

# Query strings for Empathy

«empath\*» OR «prosocial learning\*» OR «moral reasoning \*» OR « social development\*» OR «perspective taking\*» OR «cognitive empathy \*» OR « affective empathy\*» OR «prosocial behaviour \*» OR « social development\*» OR «theory of mind \*» OR « social development\*» OR « compassion \*» OR « social affect \*»

# Children

AND

«child\*» OR «toddler\*» OR student\* OR pupil\*

OR

«preschool\*» OR «pre-school\*» OR «pre-k» OR «kindergarten\*» OR «kindergartner\*» OR «ECEC» OR «ECEC» OR «ECEC» OR «daycare\*» OR «daycare\*» OR «day care\*» OR "day care\*» OR "primary school\*" OR "junior school\*" OR "elementary school\*" OR "grade school\*" OR "grammar school\*" OR "graded school\*" OR "junior school\*" OR "preparatory school\*"

AND «reading» AND «literacy» OR «literacy» OR «books» OR «story book» OR «storybook» OR «fiction» OR « literature»

«e-book\*» OR «ebook\*» OR «electronic book\*» OR «electronic storybook\*» OR «screen storybook\*» OR «interactive book\*» OR «interactive storybook\*» OR «digital book\*» OR «digital picture book\*» OR «digital storybook\*» OR «picture e-book\*» OR «app book\*» OR «e-storybook\*» OR «multimedia book\*» OR «media book\*» OR «talking book\*» OR «living book\*» OR «living storybook\*» OR «multimedia book\*» OR «animated book\*» OR «video book\*» OR «Iving book\*» OR «Iving storybook\*» OR «multimedia book\*» OR «animated book\*» OR «video book\*» OR

«story app» OR «story apps» OR «picture book app» OR «picture book apps» OR «picturebook app» OR «picturebook apps»

OR

«multimedia story» OR «multimedia stories» OR «interactive story» OR «interactive stories» OR «computer story» OR «digital story» OR «digital stories»

«CD-rom story» OR «CD-rom stories» OR «CD rom story» OR «CD rom stories» OR «DVD story» OR «DVD stories» OR «electronic console\*» OR «ec book\*»

OR

«digital reading» OR «e-literature» OR «e-reading»

# Appendix B – Quality Assessment

Gough's (2007) Weight of Evidence (WoE) criteria

A)Taking account of all quality assessment issues, can the individual study findings be trusted?

B)What is the appropriateness of the research design and analysis for addressing the aims of the individual study?

C)What is the relevance of the particular focus of the individual study for addressing its aims?

D)Taking into account the quality of execution, appropriateness of the design and relevance of focus, what is the overall weight of evidence this individual study provides to answer its research questions?

# Features evaluated with Strength of Evidence Framework (Murphy and Unthiah, 2015)

- Robustness of sample size
  - (a) High (Sample size is justified (e.g., power analyses))
  - (b) Medium (Sample size is justified but not through the use of conventional means (e.g., power analyses)
  - (c) Low (Sample size is not justified)
- Characteristics of the sample (target children aged 2-10 years)
  - (a) High (Characteristics are reported)
  - (b) Medium (Some but not all characteristics are reported)
  - (c) Low (None of the characteristics are reported)
- Definition of the outcome measure
  - (a) High (Clear definition of empathy/empathy related skills within the context of the study)
  - (b) Medium (Vague definition provided within the context of the study)
  - (c) Low (No definition provided)
- Storybook (Format and content)

- (a) High (Clearly defined and presented)
- (b) Medium (Vague definition and presentation provided)
- (c) Low (None reported)
- Intervention characteristics
  - (a) High (Clearly defined and presented)
  - (b) Medium (Vague definition and presentation provided)
  - (c) Low (None reported)