AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

ISSN: 0012-1649

https://doi.org/10.1037/dev0002005

Pretesting Promotes Theme Extraction but Not Generalization From Moral Stories in Young Children

Jyh Hsien Foo, Liwen Yu, Steven C. Pan, and Xiao Pan Ding
Department of Psychology, Faculty of Arts and Social Sciences, National University of Singapore

Although adults commonly tell moral stories to children with the aim of teaching important lessons, they often fail to extract the intended moral themes of these stories. Few interventions to date have successfully addressed this problem. The present study examined the effects of pretesting on children's learning from moral stories. We assigned sixty 5- to 6-year-olds to either a pretesting condition, where they attempted pretest questions about a to-be-learned moral before hearing a moral story, or a control condition, where they only heard the story. After hearing the story, pretested children were significantly more likely than the control group to generate the intended moral theme on open-ended questions asking about the story's lesson. They did not perform any better, however, on tasks that involved selecting the theme in a forced-choice format or applying it to novel scenarios. Hence, the present findings partially support the hypothesis that pretesting improves children's theme extraction from moral stories—that is, enhancing the ability to reproduce the correct theme but not enhancing generalization to other circumstances. Overall, this study reveals that pretesting can enhance the effectiveness of moral stories as an educational tool, at least in terms of helping children to learn important moral lessons, and presents directions for further research to investigate real-life applications of these lessons.

Public Significance Statement
The present study investigated whether pretesting—attempting to answer questions about information one has yet to learn, followed by an opportunity to learn the correct answers—helps children learn from moral stories. Five- and 6-year-olds better recalled a moral lesson after pretesting but did not generalize it to new scenarios. Hence, pretesting can enhance moral learning from stories in young children but may not guarantee broader application of those lessons.

Keywords: pretesting effect, transfer, cued recall, moral learning

Supplemental materials: https://doi.org/10.1037/dev0002005.supp

Moral stories have long been used to teach important values that are thought to shape the moral behavior of children (Bennett, 1993). These values, such as honesty and cooperation (Lee et al., 2014; Narvaez et al., 1998), are typically presented as the underlying theme or lesson of the story. Accordingly, learning from moral stories depends critically on one's ability to extract intended themes or lessons, which can lead to their successful application in real life (van den Broek et al., 2003). However, doing so can be difficult for young children, who tend to prioritize trivial story details or misinterpret themes in ways that are unintended by authors (Goldman et al., 1984; Lehr, 1988). As few interventions to date have been successful at improving children's ability to understand the intended moral themes of stories (e.g., Ding et al., 2024; Walker & Lombrozo, 2017), there is a continued need to identify pedagogical approaches that can do so. The present study investigated the effects of *pretesting*, a learning technique where practice questions are

Larisa Heiphetz Solomon served as action editor.
Jyh Hsien Foo https://orcid.org/0009-0009-0353-9830
Liwen Yu https://orcid.org/0000-0003-2414-0836
Steven C. Pan https://orcid.org/0000-0001-9080-5651
Xiao Pan Ding https://orcid.org/0000-0003-0836-9577
Jyh Hsien Foo and Liwen Yu contributed equally as joint first authors to this work. Materials and data are available on the Open Science Framework and can be accessed at https://osf.io/dsc6r/.

The authors declare that they have no conflicts of interest. This research was supported by a Faculty Development Grant from the Centre for Family and Population Research, National University of Singapore (principal investigator: Steven C. Pan; co-principal investigator: Xiao Pan Ding). The authors thank Caren Walker for sharing coding criteria.

Jyh Hsien Foo played a lead role in data curation, investigation, and writing–original draft, an equal role in formal analysis, and a supporting role in conceptualization, methodology, and writing–review and editing. Liwen Yu played a lead role in conceptualization and methodology, an equal role in formal analysis, and a supporting role in writing–original draft, writing–review and editing, investigation, and data curation. Steven C. Pan played a lead role in funding acquisition and a supporting role in conceptualization, methodology, supervision, and writing–review and editing. Xiao Pan Ding played a lead role in supervision and a supporting role in conceptualization, funding acquisition, methodology, resources, and writing–review and editing.

Correspondence concerning this article should be addressed to Steven C. Pan or Xiao Pan Ding, Department of Psychology, Faculty of Arts and Social Sciences, National University of Singapore, 9 Arts Link, Singapore City 117572, Singapore. Email: scp@nus.edu.sg or psydx@nus.edu.sg

attempted even before new information is studied (Richland et al., 2009), on enhancing children's learning from moral stories.

Children's Difficulties in Extracting the Moral Theme From Stories

Accurate extraction of a theme from a moral story requires forming an abstract mental representation of the story that connects events and ideas, without referring to specific plots or characters (Afflerbach, 1990). For example, in the popular story "The Three Little Pigs," the theme "hard work pays off" is an abstraction that goes beyond the literal elements of the plot (e.g., "building a strong house"; Lehr, 1988). Prior studies have consistently shown that theme extraction is difficult for children younger than 10 years old (e.g., Goldman et al., 1984; Jose et al., 2005; Mares & Acosta, 2008). When asked about a story's theme, children tend to face two common problems: They either misinterpret the story from an egocentric perspective, applying familiar but irrelevant values (Lehr, 1988), or fail to abstract from the story's surface details, instead merely restating concrete events (Narvaez et al., 1998).

The first problem that children frequently face in learning from moral stories is their tendency to misinterpret the main theme or lesson of the story. Researchers have attributed this issue to egocentric perspectives shaped by moral schemas and personal concerns (Lehr, 1988; Narvaez, 2002). To illustrate, Lehr (1988) found that 5- to 6-year-olds often generated incorrect themes referring to safety concerns (e.g., "do not trust strangers" as the proposed theme of "The Three Little Pigs"), likely because personal safety is a highly salient issue that is repeatedly emphasized by adult authority figures. Similarly, Brandão (1994; see also Brandão & Oakhill, 2005) found that some 4-year-olds incorrectly interpreted the theme of a new story based on their memory of familiar but unrelated fairy tales. Hence, when children encounter unfamiliar content, they may overrely on irrelevant prior knowledge to make sense of the story (Brandão & Oakhill, 2005). Consequently, children derive lessons that are unintended by the author, diminishing their ability to learn the target moral.

The second common problem to moral theme comprehension is children's inability to abstract away from the surface features of stories (Narvaez et al., 1998), which may be closely tied to their deficiencies in depth of representation. In the separate but relevant literature on problem solving, depth of representation refers to the extent to which one represents problems in terms of an overarching goal structure linking event sequences (Brown et al., 1986). There are two major reasons to believe that young children struggle to process stories at a deep level. First, they may find it difficult to ignore concrete story details. During storybook readings, children often focus more on vivid depictions of characters and actions highlighted through pictures and narration, and consequently, unobservable goals (which are key to understanding the story's theme) receive less attention (Flack & Horst, 2018). Second, making sense of a narrative's goal structure can also be difficult (van den Broek et al., 1996). Unlike adults, 4- to 6-year-old children tend to lack the sensitivity to detect links between goals, actions, and outcomes in a scenario (van den Broek et al., 1996) and, without those links, may focus on each element and associated details separately. Together, the literature shows that while children's ability to resist concrete story details and detect causal links to the goals of the characters increases with age, maturity is not reached until school age or beyond. This developmental limitation may explain why children focus on superficial details, hindering their ability to grasp a story's underlying message (Narvaez et al., 1999; van den Broek et al., 1996).

Existing Strategies to Improve Moral Theme Extraction

Given the limited theme extraction abilities of children, researchers have sought to identify techniques that can guide them toward the intended lesson. Williams et al. (2002) explored using extensive structured discussions during and after storytelling on moral theme extraction in 8- to 9-year-olds. Doing so improved children's ability to recognize the instructed morals (e.g., perseverance) when they appeared in new stories, but did not enhance the application of those skills to novel stories that required generating unfamiliar themes (i.e., cooperation). In another study, Mares (2006) examined the effects of repetition on 6- to 8-year-old children's comprehension of video stories. Although repeated exposure deepened their evaluation of the characters, children were still unable to identify the main lesson. Thus, not only might time constraints limit the feasibility of repeated exposure (especially in classroom settings), but it also appeared to be not particularly effective.

In contrast, interventions that require children to generate responses to questions may be more promising. After reading a story, Carnine et al. (1982) had 9- to 12-year-olds (children in Grades 4 to 6) answer facilitative questions that required them to reflect and explain character actions and story outcomes. This technique helped children form inferences about character goals, which is often necessary to understand a story's underlying theme. Similarly, Walker and Lombrozo (2017) prompted 5- to 6-year-old children to explain key story events by posing "why" questions midway and at the end of the story. Compared to children who only restated if an event had occurred and those directly given the answers, children who generated their own explanations were better able to extract the moral lesson and identify it in other contexts. Together, these studies suggest that even 5-year-old children can understand a moral theme with appropriate guidance, particularly from questions that require reflecting on a story's abstract concepts (e.g., goals or reasons for events). In a similar vein, asking children questions even before they are exposed to a story, as in the case of pretesting, might have comparable effects.

Pretesting as a Potential Strategy to Enhance Learning From Moral Stories

Answering practice questions about to-be-learned information even before studying, or pretesting, is a promising technique for improving memory of the to-be-learned content as evidenced by better performance on subsequent tests (for a review, see Pan & Carpenter, 2023). Typically, studies of pretesting compare two conditions: a pretesting condition that engages in pretesting prior to engaging with target learning materials (e.g., a text passage or lecture), as well as a control condition that does not engage in pretesting and instead performs completely unrelated activities (e.g., solving math problems) or nontesting activities (e.g., reading statements) before engaging with the target learning materials. Learning outcomes are assessed on a subsequent test. On that test, an advantage of pretesting over the non-pretesting control condition—which is known

as the *pretesting effect*—is commonly observed. Past studies have demonstrated the pretesting effect for various types of materials, including expository texts (e.g., Pressley et al., 1990), video lectures (e.g., Carpenter & Toftness, 2017), and foreign language vocabulary (e.g., Potts & Shanks, 2014), but not yet for moral story learning. Researchers have proposed various theoretical mechanisms to explain the pretesting effect, including semantic activation, where pretest questions activate prior knowledge schemas that facilitate the encoding of study material (Kornell et al., 2009); changes in intentional learning behaviors, such as increased attention to critical parts of the material in an effort to reduce knowledge gaps (Carpenter & Toftness, 2017; McCrudden & Schraw, 2007); heightened curiosity, motivation, and increased focus (Geller et al., 2017; Potts & Shanks, 2014); and the errors generated on pretests serving as mediators to facilitate the recall of correct information (Kornell et al., 2009), among others.

Implementations of pretesting vary in the literature, with some common approaches and notable exceptions. For instance, in the context of learning from text or video, a typical approach has been to simply ask participants to attempt a series of pretest questions before reading the text, in which they can discover the correct answers, with no other intervening activities (e.g., Carpenter & Toftness, 2017; Richland et al., 2009). A few studies, however, have provided correct answer feedback after pretest attempts and before target materials are presented (e.g., Pan & Sana, 2021). On the other hand, in the case of pretesting and learning simple verbal materials, such as foreign language vocabulary, it is common for participants to receive immediate correct answer feedback after each pretest attempt (e.g., Potts & Shanks, 2014), with that feedback constituting the target learning materials. In a few studies of pretesting and text or video learning, participants also recalled pretest questions prior to encountering the target materials (e.g., St. Hilaire et al., 2019), ostensibly to check whether they were paying attention.

A small body of evidence suggests that the benefits of pretesting can be extended to children (see also Brod, 2021). In the case of pretesting with expository texts and similar materials, studies have demonstrated pretesting effects relative to a non-pretesting condition (Shanahan, 1983) or a restudy condition (de Lima & Jaeger, 2020) among children as young as elementary school age. Similarly, children are more likely to remember general knowledge facts or cue-target associations after initially generating error-prone predictions about them (Breitwieser & Brod, 2021; Carneiro et al., 2018). Not all studies of pretesting and young learners have shown pretesting effects, however (e.g., McDaniel et al., 2011), with methodological details being a possible culprit (for discussion, see Pan & Carpenter, 2023).

Crucially, researchers have yet to examine whether pretesting may benefit children's learning of abstract concepts such as morals. Drawing on the theoretical explanations for the pretesting effects observed in past studies, the present study proposes that pretesting may be able to improve children's theme extraction from moral stories. If so, then pretesting may facilitate children's ability to generalize the learned moral themes to reasoning about real-life scenarios.

How Pretesting Might Affect Moral Theme Extraction and Generalization

One way in which pretesting may be useful for learning from moral stories is by helping children overcome their egocentric tendencies. By one account, pretesting activates schemas relevant to the target material (Kornell et al., 2009), which in the case of moral stories may involve past experiences or concepts that are specifically related to the moral (Narvaez, 2002). The attempt to answer a pretest question may activate topically relevant knowledge in semantic memory, with the activated concepts later used to process the target story in a focused direction (Pressley et al., 1992; Richland et al., 2009). Absent such activation, children may instead misinterpret a moral story based on personally important or familiar values (Lehr, 1988). Pretesting may help children to organize prior knowledge and orient their thinking toward the intended themes.

By another account, pretesting may address children's difficulties in abstracting away from the surface value of stories. It has been suggested that pretesting can promote deeper processing of to-be-learned material when it is later encountered (Murphy et al., 2023; Rickards, 1976). To search the target material for answers to conceptually focused pretest questions (such as questions related to an abstract moral), one often needs to identify connections and integrate relevant information from different passage sections (Memory, 1983; St. Hilaire et al., 2019). The need to form a relational structure might prevent young children from prioritizing individual story units and their superficial details. Accordingly, pretesting may encourage children to organize both pretested and other to-be-learned information around a core idea instead of focusing on separate concrete events.

Besides addressing the two main obstacles that children face in understanding a story's moral, pretesting may also generate curiosity about the answers, resulting in improved attention and more careful processing of the target material (Geller et al., 2017; Pan et al., 2020). If so, then pretesting may help children overcome their vulnerability to distractions and struggles with the perceptual load during storytelling sessions (Flack & Horst, 2018; Vaahtoranta et al., 2018). Increased attention may also be directed toward critical information, aiding in resolving errors rather than being distracted by unimportant superficial details (McCrudden & Schraw, 2007). All these processes may yield better learning from moral stories.

Since the efficacy of moral education tools ultimately lies in their ability to guide everyday behaviors (Schoppmann et al., 2023; Strouse et al., 2018), the present study also investigated whether pretesting enables children to generalize a story's lesson to realistic contexts. Children's understanding of stories is often tied to literal aspects, making it difficult to make connections with everyday experiences (Strouse et al., 2018). If an abstract representation of a story's structure is formed, that can help them recognize analogies to new experiences, improving transfer (Brown et al., 1986; Mar & Oatley, 2008). Moreover, successful comprehension might promote the internalization of related moral values, thereby increasing the likelihood of ethical decision making in future situations (Kochanska, 2002). Hence, if pretesting can improve children's theme extraction from moral stories, then it may also enable them to better transfer moral lessons to reasoning about real-life scenarios.

The Present Study

We examined whether pretesting can improve 5- to 6-year-old children's learning from moral stories. Although most studies of pretesting have been conducted with adults, and a few studies have involved elementary school children, it is not yet clear whether

preschool children may also be receptive to the benefits of the technique. We focused on 5- to 6-year-olds for three reasons. First, children in this age range can form guesses based on existing schemas and process corrective feedback (Berner et al., 2022; Carneiro et al., 2018). Second, children of this age group do not spontaneously extract moral themes but can do so with guidance (Walker & Lombrozo, 2017). Finally, previous studies on children's narrative comprehension intervention have examined a similar age group (e.g., Lever & Sénéchal, 2011), and hence, the present study could be compared to such prior work.

In the present study, we randomly assigned 5- and 6-year-olds to a pretesting or control condition. Before presenting a moral story, we asked those in the pretesting condition scenario-based pretest questions that were related to the to-be-learned moral. Similar to some prior studies of pretesting (Rickards, 1976; St. Hilaire et al., 2019), we constructed conceptually focused pretest questions (e.g., "why" questions) in which the search for answers would require integration between different sections of the target story instead of directing attention to any one part of the text. After attempting the questions, children in the pretesting condition received feedback that might highlight knowledge gaps and stimulate curiosity about the lesson (Brod, 2021). They were also told that the upcoming story would aid them in answering the questions correctly in the future, as such instructions can motivate learners to remember the pretest questions and actively search for answers (St. Hilaire & Carpenter, 2020; St. Hilaire et al., 2019). Children in the control condition, on the other hand, completed an unrelated drawing task before hearing the story.

We hypothesized that pretesting (a) improves children's theme extraction from moral stories and (b) facilitates children's generalization of a story's lesson to reasoning about novel, real-life contexts. To assess learning outcomes, we adapted three measures from Walker and Lombrozo (2017). We assessed theme extraction using two tasks: an open response task asking about "the most important lesson" learned, as well as a theme selection task where children selected one of three possible answers to the same question. We measured generalization ability with three open-ended transfer questions that required applying the moral lesson to solve problems in hypothetical real-world scenarios.

Method

Participants

We aimed to conclude participant recruitment upon reaching valid data from 60 children (30 children per condition). This stopping rule was preregistered and modeled after Walker and Lombrozo (2017), who included 24–32 participants per condition in a between-participants design that examined the effect of explanations on moral learning. Formal data analysis included 60 typically developing 5- and 6-year-olds (M_{age} = 70.90 months, SD_{age} = 6.36 months; 30 males), with 30 participants randomly assigned to each condition (pretesting or control). We recruited an additional 13 children but excluded them because they were unresponsive (n = 8) or disruptive (n = 4) or their parents interrupted the procedure (n = 1). Participants were either recruited and tested at two preschools in Singapore (n = 16) or were recruited through social media advertisements and were tested in a university laboratory (n = 44). We conducted a sensitivity analysis using GPower 3.1 (Faul et al., 2009), which indicated that our sample of 60 participants provided 80% power to detect a small to medium effect size ($f^2 = 0.14$) in a hierarchical multiple regression with one tested predictor (i.e., learning condition) at a significance level of $\alpha = .05$.

The participants were of various ethnicities (90.0% Chinese, 5.0% Indian, 1.7% multiracial, 3.3% others) and could all understand and speak English. Prior to their participation, we obtained informed parental consent and children's verbal assent. This study was approved by the National University of Singapore Institutional Review Board (Protocol Number: NUS-IRB-2023-365).

Materials

We used two moral stories from Walker and Lombrozo (2017), with each story assigned to half of the participants in each condition, following their established approach. "Tall People" was a story about *being nice to others even if they are different*. In this story, the tall people in a town initially made fun of a man for being short but eventually learned to be nice to him. "The Queen's Painting" emphasized the importance of *working together to achieve greater things*. This story described three families who each wanted to use only their favorite color for a painting but later learned that they could make beautiful artwork by combining their colors and working together. Each story had 14 pages of printed illustrations.

Additionally, we designed pretest questions corresponding to each story. Specifically, we created three short scenarios that had similar plotlines to each respective story, each ending in two "direct" pretest questions that required participants to interpret the scenario as well as a "reasoning" pretest question that required participants to explain their moral reasoning. For the "Tall People" story, all scenarios described someone being excluded from a group for being different, whereas for "The Queen's Painting" story, all scenarios described people who faced challenges when they worked on a task alone.

The first scenario always involved similar characters and settings as the upcoming story, while the following two involved different ones. For example, the first "Tall People" pretest scenario referred to the exclusion of a short person:

> There is a person who is very different because he is very short. One day, this group of people met with this person. They begin to laugh at him and didn't want to play with him because he is very short. How would that person feel? Is it a nice thing for this group of people to do that? Why?

The next two pretest scenarios were about excluding a person who spoke a different language or only had four fingers.

Within each scenario, questions were centered on the to-be-learned moral. In the "Tall People" example, we asked children how the excluded person would feel and whether it was nice of the group to exclude him (both as "direct" pretest questions) and to explain why or why not (as a "reasoning" pretest question). Each child in the experimental condition responded to all three scenarios in the same order.

It should be noted that we constructed pretest questions/scenarios that involved different content and required qualitatively different responses (i.e., value judgments) than those used in the majority of the pretesting literature. Most prior studies using pretest questions have asked participants to make guesses about facts (e.g., Kornell et

al., 2009) or concepts (e.g., Richland et al., 2009) that do not have an explicit moral dimension. Moreover, in such studies, it is usually quite difficult to successfully guess the factually correct answer to such questions, whereas with the questions in the present research, it is possible, even likely, to answer the "direct" pretest questions based on careful attention to (and thoughtful reasoning about) the presented scenario. Further, there was not necessarily an unambiguously factually correct answer to such questions; rather, there was an "intended" answer that best matched the to-be-learned moral (it should be noted that reasoning about moral scenarios is often inherently ambiguous and can yield several answers that might be judged to be morally defensible). As previously noted, however, the "reasoning" pretest questions were more difficult and resembled conceptually focused questions in some prior work (e.g., Rickards, 1976; St. Hilaire et al., 2019) by requiring participants to integrate different sections of the presented scenario.

Tasks

Participants completed three tasks that assessed moral learning (Open Response, Theme Selection, and Transfer Tasks), as well as a separate set of working memory capacity and language tasks.

Open Response Task

After hearing the story, all children were asked about the most important lesson they learned from it. As in Walker and Lombrozo (2017), we categorized children's responses as either lesson-based, content-based, or irrelevant. Lesson-based responses signified the ability to extract the relevant theme in general terms (e.g., accepting others; working together), while content-based responses focused exclusively on superficial details (e.g., being short; mixing colors). Irrelevant responses referred to any unintended themes or details that did not answer the question. We obtained a binary open response score by assigning a 1 to lesson-based responses and a 0 for the other categories.

Theme Selection Task

Children were given three choices and had to select the most important lesson that they learned in the story. One option referred to the intended lesson (e.g., "it doesn't matter if people are different, we can still be nice to each other and have fun together"), the second option referred to specific content from the story (e.g., "the people from Talleg laughed at Jocko and didn't want to play with him at first"), and the last option was an unintended lesson (a subtheme that was not the most important lesson; e.g., "playing games is fun"). The experimenter presented each option on a card that was read aloud. The order in which the three options were presented was counterbalanced to create two possible sequences that participants were randomly assigned to. We awarded a score of 1 if children selected the intended lesson-based choice and a 0 if they chose either the content-based or unintended lesson option.

Transfer Task

We used three scenario-based questions to assess each child's ability to generalize the story's moral to real-life contexts. The order of presentation was counterbalanced in the same way as in the theme selection task. Each scenario was accompanied by printed images and introduced as a real event experienced by the experimenter. For all scenarios, children were asked "what should [character(s)] do?" and why. We awarded a score of 1 only if the child's solution was an application of the story's lesson (e.g., they should let the girl join even if she looks different), and coded all other responses as 0. We summed the scores for the three questions into a total transfer score (0–3).

Working Memory and Language Tasks

Children's understanding of a story is associated with their working memory capacity (van den Broek et al., 2005; Williams et al., 2002) and language ability (Mares, 2006; Pelletier & Beatty, 2015). To rule out possible preexisting differences between conditions that could affect the dependent measures, we measured and controlled for both factors.

The Backward Digit Span (BDS) task was used as a measure of working memory (Bialystok, 2010; Carlson et al., 2002; Davis & Pratt, 1995). On each test trial, children listened to a string of numbers read aloud once by the experimenter (from the National Longitudinal Survey of Youth, NLSY79 Child Surveys, 1974) and had to recall the numbers in the reverse order. The length of the longest digit string that could be successfully recalled was taken as the BDS score. For children who failed even the two-digit test trials, a score of 1 was assigned instead. Hence, the possible range for BDS scores was 1 to 8.

The English Peabody Picture Vocabulary Test (PPVT), Fourth Edition (Dunn & Dunn, 2007), was used as a measure of language ability. The task was administered according to standardized procedures, and age-corrected standard scores were used in the analyses.

Procedure

We tested children individually, in English, in a preschool classroom or university laboratory room. In the pretesting (experimental) condition, children first attempted to answer questions about three pretest scenarios that were accompanied by illustrations. They received general feedback for each question to inform them if their answers were inconsistent with the intended moral lesson ("thanks for sharing your ideas—however, my answer is different from yours"), consistent ("those are good answers"), or lacking in justification despite being aligned with the lesson ("as for the reasons, we can think more about it"). Regardless of their answers and the feedback provided, the experimenter told the children in this condition that they could learn how to answer the questions from an upcoming story. In the control condition, children were given 5 min to draw and did not answer any questions.

All children were subsequently introduced to the moral story. They were instructed to listen carefully as the story would teach them an important lesson. To check their understanding, all children were asked to repeat the instructions. Those in the experimental condition were additionally reminded that this lesson would help them to answer the pretest questions that they had attempted before and were required to recall them ("do you remember what the questions are about?"). The experimenter laid illustrations of the pretest questions on the table to aid children's responses and provided feedback until they could recall all three questions correctly.

Children in both conditions then listened to the story that was read aloud by the experimenter.

Immediately after hearing the story, children were asked four memory questions regarding story events to ensure that they were paying attention. They replied "yes" or "no" to these questions and received feedback that reiterated or clarified the event. Next, children were taught the definition of a "lesson" as the message that the person who created the story wanted us to learn from it. This brief training attempted to reduce any ambiguity surrounding the word and to facilitate thinking about the lesson or theme of the story that they just heard (Walker & Lombrozo, 2017). Children then completed the dependent measures in the following order—the open response task, the theme selection task, followed by the transfer task.

After the main tasks, children completed the working memory task (Backward Digit Span), followed by the language task (Peabody Picture Vocabulary Test), and were then debriefed and reimbursed with a gift.

Data Coding and Analysis

Two researchers independently coded the open-ended responses of all participants. For the open response task, the two coders agreed on 95% of the scores. The intraclass correlation coefficient (ICC) was .90. The interrater reliability for the transfer task was even higher, as the coders agreed on more than 99% of the scores (ICC = .98). Any disagreements were resolved through a discussion between the two coders and other members of the same research lab.

The analysis plan began with computing descriptive statistics for performance in the pretested condition on the different pretest question types. Next, as in the preregistered analysis plan, we conducted a hierarchical logistic regression with the binary open response score as the outcome variable and condition as a categorical predictor, with age in months, working memory (BDS score), and language ability (PPVT score) as planned covariates. This analysis was repeated with the theme selection score as the outcome variable. We also performed a hierarchical linear regression on the overall transfer score, controlling for the same covariates.

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures, and we follow Journal Article Reporting Standards (Appelbaum et al., 2018). All data, analysis code, and materials are available at https://osf.io/dsc6r/ (Foo et al., 2025). We analyzed the data using R, Version 4.3.1 (R Core Team, 2023), and preregistered the study design, hypotheses, and analysis plan at AsPredicted.org (https://aspredicted.org/qqyt-3s39.pdf).

Results

Pretest Question Performance

We computed descriptive statistics on both pretest question types, collapsed across versions of each question. For the "direct" pretest questions—such as "Is it nice of the group to exclude him?"—the mean percentage of responses that aligned with the intended moral lesson (i.e., "no" to the aforementioned question) was 96% (SD = 15%). For the "reasoning" pretest questions—such as "Why is it (not) nice?"—this percentage was 43% (SD = 39%). It is notable that these performance levels are much higher than are typically observed in the pretesting literature (Pan & Carpenter, 2023). An analysis involving performance across scenarios and relating that performance to open response scores is presented in the online Supplemental Materials.

Task Performance

Descriptive statistics and response patterns for all three tasks (open response, theme selection, and the transfer task scores) are depicted in Figure 1.

Preliminary Analyses

Preliminary analyses of task performance revealed no significant gender differences in performance on the dependent measures. Hence, data were collapsed across gender for subsequent analyses. Simple logistic regression analyses revealed that story version (0 = The Queen's Painting, 1 = Tall People) had a significant effect on the binary open response scores ($B = 2.16$, Wald = 12.12, $p < .001$, $OR = 8.64$) and theme selection scores ($B = 1.54$, Wald = 7.69, $p = .006$, $OR = 4.67$). In both tasks, children who heard the "Tall People" story performed better than those who heard "The Queen's Painting." Story version did not significantly predict transfer scores. Thus, in addition to the preregistered covariates of age in months, BDS, and PPVT scores, we also controlled for story version in the analyses of open response and theme selection scores when comparing between the pretesting and control conditions.

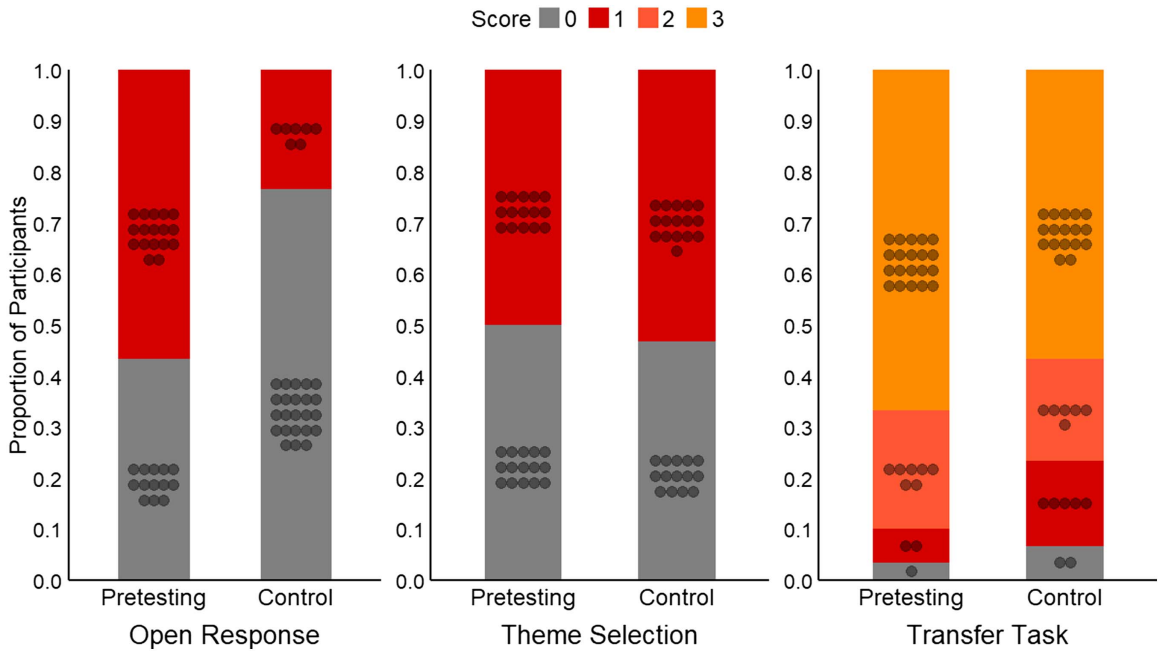
Testing location (0 = preschool, 1 = lab) had a significant effect on the transfer scores in a simple linear regression model ($\beta = 0.32$, $t = 2.55$, $p = .014$), indicating that children who were tested in the lab performed better than those tested in the preschools on the transfer task. We did not find this effect for the other dependent measures. Hence, in addition to the three preregistered covariates, we also controlled for testing location in the analysis of the transfer scores. Descriptive statistics by condition are presented in Table 1.

Theme Extraction

To investigate the hypothesis that pretesting improves children's theme extraction from moral stories, we examined the performance on the open response and theme selection tasks.

Open Response Score. We conducted a hierarchical logistic regression on the binary score from the open response task (0 = no abstraction, 1 = lesson abstraction). Age in months, BDS scores, PPVT scores, and story version (0 = The Queen's Painting, 1 = Tall People) were entered in the first step as control variables. Condition was added in the second step. The first model was significant, $\chi^2(4, N = 60) = 22.76$, Nagelkerke $R^2 = .43$, $p < .001$. PPVT scores and story version significantly predicted open response scores ($B = 0.08$, Wald = 4.74, $p = .029$, $OR = 1.09$; $B = 2.83$, Wald = 14.00, $p < .001$, $OR = 16.89$). The results indicated that children were more likely to generate a relevant lesson-based response as language ability increased and that their performance was better for "Tall People" than for "The Queen's Painting." The second model was also significant, $\chi^2(5, N = 60) = 35.02$, Nagelkerke $R^2 = .60$, $p < .001$. The likelihood ratio test showed that the second model could explain significantly more variance in open response scores than the

Figure 1
Task Performance



Note. Dots represent the individual participant scores. Scores on the open response and theme selection were 0 or 1, whereas the transfer task scores ranged from 0 to 3. A higher score indicates more successful responding. See the online article for the color version of this figure.

first model, $\Delta\chi^2(1, N = 60) = 12.26$, ΔNagelkerke $R^2 = .17$, $p < .001$. While the contributions of PPVT scores and story version remained significant, the contribution of BDS scores reached marginal significance in this model ($B = 1.20$, Wald = 3.45, $p = .063$, $OR = 3.32$). Table 2 provides details of the analyses.

Importantly, condition was found to significantly predict open response scores over and above the other contributions ($B = 2.95$, Wald = 7.97, $p = .005$, $OR = 19.14$). Specifically, children in the pretesting condition were more likely than those in the control condition to generate the intended abstract theme in the open response task. This result indicates better learning from the moral stories following pretesting.

Theme Selection Score. We performed a hierarchical logistic regression with theme selection scores (0 = incorrect choice, 1 = correct lesson-based choice) as the outcome variable. Age, BDS scores, PPVT scores, and story version were entered in the first step. Condition was entered in the second step. The first model was significant, $\chi^2(4, N = 60) = 13.55$, Nagelkerke $R^2 = .27$, $p = .009$. Story version was a significant predictor ($B = 1.72$, Wald = 7.87, $p = .005$, $OR = 5.56$), indicating that performance on the theme selection task was better for "Tall People" than for "The Queen's Painting." The effect of age was also marginally significant ($B = 0.10$, Wald = 3.54, $p = .060$, $OR = 1.10$). While the second model was significant, $\chi^2(5, N = 60) = 13.62$, Nagelkerke $R^2 = .27$, $p = .018$, the likelihood ratio test showed that the second model could not explain significantly more variance in theme selection scores than the first model, $\Delta\chi^2(1, N = 60) = 0.07$, ΔNagelkerke $R^2 = .001$, $p = .796$. Hence, while story version remained as a significant predictor and age as a marginally significant predictor, there was no significant effect of condition over and above the other contributions in the model ($B = -0.15$, Wald = 0.07, $p = .796$, $OR = 0.86$). Thus, children in the pretesting and control conditions did not differ in their tendency to select the correct lesson-based choice in the theme selection task. Table 3 provides details of the analysis.

Table 1
Descriptive Statistics of Variables by Condition

Measure	Pretesting ($n = 30$)		Control ($n = 30$)	
	M	*SD*	*M*	*SD*
Age (months)	70.83	5.78	70.97	6.99
Open response score (0/1)	0.57	0.50	0.23	0.43
Theme selection score (0/1)	0.50	0.51	0.53	0.51
Transfer task score (0–3)	2.53	0.78	2.27	0.98
BDS score (1–8)	3.03	0.93	3.10	0.80
PPVT score (standardized)	108.43	8.46	106.27	11.27

Note. Possible ranges of scores are in parentheses. BDS = Backward Digit Span; PPVT = Peabody Picture Vocabulary Test.

Generalization to Novel Contexts

To examine the second hypothesis that pretesting facilitates children's generalization of the story's lesson to novel real-life contexts, we analyzed the overall score on the transfer task (0–3) as the outcome variable in a hierarchical linear regression. Age, BDS scores, PPVT scores, and testing location (0 = preschool, 1 = lab) were entered in the first step. Condition was entered in the second

Table 2
Hierarchical Logistic Regression of Open Response Scores

Variable	Step					Model		
	B	*SE*	Wald	*p*	*OR*	χ^2	Nagelkerke R^2	*p*
Step 1						22.76***	.43	<.001
Age (months)	0.03	0.06	0.33	.564	1.03			
BDS score	0.43	0.45	0.89	.346	1.53			
PPVT score	0.08*	0.04	4.74	.029	1.09			
Story (Tall People)	2.83***	0.76	14.00	<.001	16.89			
Step 2						35.02***	.60	<.001
Age (months)	0.02	0.07	0.08	.783	1.02			
BDS score	1.20†	0.65	3.45	.063	3.32			
PPVT score	0.11*	0.05	4.69	.030	1.11			
Story (Tall People)	4.06***	1.15	12.47	<.001	57.83			
Condition (pretesting)	2.95**	1.05	7.97	.005	19.14			

Note. For story, The Queen's Painting is the reference category. For condition, control condition is the reference category. *SE* = standard error; BDS = Backward Digit Span; PPVT = Peabody Picture Vocabulary Test.
† Marginally significant ($.05 < p < .10$). * $p < .05$. ** $p < .01$. *** $p < .001$.

step. The first model was not significant, $R^2 = .12$, $F(4, 55) = 1.85$, $p = .133$. However, among the predictors entered in this step, testing location was a significant predictor ($\beta = 0.28$, $t = 2.03$, $p = .048$), indicating that children who were tested in the lab scored higher on the transfer task than those tested in the preschools. The second step did not significantly improve the model's predictive ability, $\Delta R^2 = .03$, $\Delta F(1, 54) = 1.74$, $p = .192$. There was no significant effect of condition over and above the other contributions in the model ($\beta = 0.17$, $t = 1.32$, $p = .192$). The results suggest that scores on the transfer task did not differ between children in the pretesting and control conditions. Table 4 provides details of the analysis.

Discussion

We investigated the effects of pretesting on 5- to 6-year-old children's learning from moral stories. We found partial support for the hypothesis that pretesting can improve children's ability to extract a story's moral theme. Compared to children in the control condition, children who attempted pretest questions were more likely to provide an abstract lesson-based response when asked to generate the story's theme in an open-ended format. However, pretesting did not improve children's ability to identify the intended lesson in a forced-choice theme selection task. Moreover, the results did not support the second hypothesis that pretesting can improve children's ability to generalize the story's lesson to novel, real-life contexts.

How Pretesting Impacted Learning From Moral Stories

Whereas prior studies have shown that children often fail to extract the theme of a story (Goldman et al., 1984; Jose et al., 2005; Mares & Acosta, 2008), our results reveal that attempting to answer questions related to the intended lesson even before hearing the story can help children as young as 5 and 6 years old overcome their difficulties. When generating an open-ended response to the story's lesson, children in the pretesting condition were more likely than

Table 3
Hierarchical Logistic Regression of Theme Selection Scores

Variable	Step					Model		
	B	*SE*	Wald	*p*	*OR*	χ^2	Nagelkerke R^2	*p*
Step 1						13.55**	.27	.009
Age (months)	0.10†	0.05	3.54	.060	1.10			
BDS score	0.09	0.39	0.05	.822	1.09			
PPVT score	−0.01	0.03	0.19	.665	0.99			
Story (Tall People)	1.72**	0.61	7.87	.005	5.56			
Step 2						13.62*	.27	.018
Age (months)	0.10†	0.05	3.57	.059	1.10			
BDS score	0.08	0.39	0.04	.832	1.09			
PPVT score	−0.01	0.03	0.16	.685	0.99			
Story (Tall People)	1.72**	0.61	7.90	.005	5.60			
Condition (pretesting)	−0.15	0.59	0.07	.796	0.86			

Note. For story, The Queen's Painting is the reference category. For condition, control condition is the reference category. *SE* = standard error; BDS = Backward Digit Span; PPVT = Peabody Picture Vocabulary Test.
† Marginally significant ($.05 < p < .10$). * $p < .05$. ** $p < .01$.

Table 4
Hierarchical Linear Regression of Transfer Scores

Variable	Step				Model				
	β	*SE*	*t*	*p*	*df*	*F*	R^2	ΔR^2	*p*
Step 1					4, 55	1.85	.12	.12	.133
Age (months)	0.14	0.02	1.02	.315					
BDS score	−0.06	0.14	−0.43	.670					
PPVT score	0.08	0.01	0.54	.594					
Testing location (lab)	0.28*	0.27	2.03	.048					
Step 2					5, 54	1.85	.15	.03	.120
Age (months)	0.13	0.02	0.95	.347					
BDS score	−0.05	0.14	−0.33	.742					
PPVT score	0.05	0.01	0.32	.753					
Testing location (lab)	0.30*	0.27	2.19	.033					
Condition (pretesting)	0.17	0.22	1.32	.192					

Note. For testing location, preschool is the reference category. For condition, control condition is the reference category. *SE* = standard error; BDS = Backward Digit Span; PPVT = Peabody Picture Vocabulary Test.
* $p < .05$.

those in the control condition to generate the intended theme, which suggests that they may have privileged an abstract representation of the story over superficial details. This pattern potentially occurred because relevant schemas about the moral lesson were activated during pretesting (Kornell et al., 2009), helping children to organize prior knowledge and subsequently process the story in the intended way. If so, this would align with the possibility that having to search for answers to conceptually oriented pretest questions prompts deeper processing and integration between different story units, diverting focus away from individual surface details that are unimportant on their own (Memory, 1983; St. Hilaire et al., 2019). Moreover, as we asked children multiple pretest questions related to a single theme, they may have tried to identify similar elements among the questions, which could have prompted them to adopt an abstract thinking mode before encountering the story. In our view, these reasons likely explain why pretesting enabled children to extract and state the intended lesson in an open response task.

However, while we observed the benefits of pretesting when children verbalized the theme in an open-response format, children who were pretested were not more likely than the control group to identify the intended lesson in the theme selection task. One possible explanation could be that when shown three statements, all of which made sense in the story's context, children may have been influenced to select the safest option, that is, reiterating a concrete event that they were sure had occurred. Even if they had extracted the intended theme and acknowledged its significance among the options, the nature of the task wherein only one statement could be "correct" may have lowered their confidence in the abstract themes that they recognized but did not directly observe in the story. Moreover, children may experience choice overload (Scheibehenne et al., 2010), leading them to select the most distinct option, one that describes specific details in contrast to two other options with more abstract messages. Hence, the notion that there is strictly one answer may have influenced children in the pretesting condition to continue to select the safer content-based option even if they had successfully recognized the theme. A related possibility involves the inherent ambiguity of moral scenarios, many of which could have more than one "correct" answer. That ambiguity may have altered the parameters within which children made their responses.

Nonetheless, although the two tasks provided inconsistent results for the hypothesis that pretesting facilitates moral theme extraction, the significant finding in the open response task is still noteworthy given its difficulty. Specifically, *generating* the story's theme in one's own words has been proposed to be more challenging than *recognizing* the theme from a list of options (Narvaez et al., 1998). This perspective is supported by our observation that performance on the theme selection task was higher than on the open response task in the control condition. According to Lehr (1990), even if children can select the correct theme at a young age, they may still fail to generate a theme in their own words as doing so requires a complex integration of prior knowledge and story events into a clear coherent statement. Yet, the improved performance on the open response task suggests that pretesting could have helped children to organize topically relevant knowledge and the story material into a meaningful structure (Rickards, 1976), from which the lesson could be more easily extracted from and condensed into a single abstract statement. Hence, the significant effect of pretesting in the more difficult open response task underscores the benefits of pretesting on children's moral theme extraction.

It should also be noted that the pretesting procedure that we used was more multifaceted than that commonly employed in studies of pretesting and learning from text or video with adults. Specifically, not only did we provide detailed corrective feedback, but we also had participants recall the pretest questions prior to the moral stories and, moreover, told participants that the stories would help them better answer the pretest questions. Indeed, these design features are analogous to those used in some prior studies of pretesting with adults (e.g., Pan & Sana, 2021; St. Hilaire et al., 2019) and, at least in the case of corrective feedback, appear to facilitate the pretesting effect in adult learners. Those facilitation strategies may have been particularly beneficial for young children, who are likely to require additional support due to their developing memory and cognitive abilities. Future research is needed to determine whether such support is indispensable—specifically, whether the absence of feedback and other scaffolding would diminish the pretesting effect in young learners (relatedly, preliminary evidence suggests that younger learners do indeed require additional support at least in the form of correct answer feedback, to benefit from pretesting; Yu et al., 2025).

Another point of departure from the present versus prior research is that moral learning is qualitatively different from factual or other forms of conceptual learning that have been characteristic of the pretesting literature. In the latter cases, there often are unambiguously right or wrong answers to pretest questions and the to-be-learned information tends to be quite specific. Moreover, major theoretical accounts of the pretesting effect—such as error correction—are relatively straightforward to apply to such scenarios (i.e., the learner makes an incorrect guess to a pretest question, then discovers the precise correct answer afterward, with that answer overriding the incorrect guess in memory), but less so for the case of moral learning. It is possible that in the case of moral stories, pretesting engages different cognitive mechanisms and operates differently than for factual or other types of conceptual materials.

Generalization of Moral Learning

We hypothesized that pretesting would facilitate children's generalization of the moral theme to novel, real-life scenarios. Although the results were in the predicted direction wherein the pretesting group numerically outperformed the control group on the transfer task, this difference was not significant. There are a few possible explanations for this finding. On one hand, it is possible that pretesting does not actually have an effect on children's ability to generalize a story's moral to other contexts. It may be that pretesting has specific effects that are limited to helping children form an abstract representation of a target scenario (i.e., the moral story), or of extracting a specific theme from that scenario, but not of others. Indirect evidence for this claim comes from studies demonstrating that pretesting does not always facilitate performance on questions that were not formerly tested (Carpenter et al., 2023), which in the present study would refer to the novel transfer questions. In order for transfer to occur, however, forming an abstract mental representation of the novel scenario is also necessary (Fisch et al., 2005). Hence, while pretesting may have helped children to abstract the story itself (or at least extract the relevant theme), they could still have been distracted by the complexities and surface details of a real-life scenario, hindering them from noticing that its structure and solution are similar to that of the story.

While such an explanation should not be discounted, the descriptive statistics instead seem to suggest that children generally did not face difficulties proposing lesson-based solutions to the transfer questions. A ceiling effect may have precluded the pretesting manipulation from increasing children's transfer ability, as even the control group was found to perform well by providing responses aligned with the intended lesson on the transfer task ($M_{\text{control}} = 2.27$ out of 3, or 76%; $M_{\text{pretesting}} = 2.53$, or 84%). This effect could be due to the procedure itself, where all children were explicitly instructed to reflect on the story's moral in the prior theme extraction tasks. As such, it could have been easy for them to acknowledge that the story's lesson and solutions were relevant to the following transfer scenarios (Russell & Cain, 2022). Furthermore, children only needed to provide solutions to problems faced by other individuals. In contrast, a task involving decisions about their own behaviors would be expected to be more challenging for children, as their reasoning and intentions do not necessarily translate into actual actions following moral stories (e.g., Du & Hao, 2018). Such a task could be used as a more sensitive measure of children's application of the lesson to novel, real-life contexts. Thus, while the present study did not find significant transfer results, further investigations are needed to test the effects of pretesting on children's ability to generalize moral learning.

Study Limitations

There are several limitations to this study that should be acknowledged. First, the two stories used unintentionally varied in difficulty, with scores on the open response and theme selection tasks being significantly higher for the "Tall People" story compared to "The Queen's Painting." This disparity could be because children who were read a story about *being nice to others even if they are different* may have found the events more familiar than those told a story about *working together to achieve greater things*. Although this study did not assess participants' familiarity with a moral, children would most likely have been formerly taught the need to always be respectful and accepting toward others—a lesson that is often deemed as morally relevant and consistent in their everyday lives. In contrast, there may have been inconsistencies in prior experiences in terms of whether individuals should work together or alone (e.g., when teachers highlight the importance of attempting a task independently), especially within the cultural context of Singapore where schools often emphasize an integration of both independent and interdependent values (Wu & Tan, 2021). The present findings coincide with suggestions from past research that story themes may differ in their ease of extraction due to certain factors, such as familiarity (Afflerbach, 1990; Narvaez, 2002) or the concreteness of the behaviors dictated by the lesson (e.g., sharing vs. being honest; Fisch et al., 2001).

Another methodological limitation pertains to environmental factors. Children who were tested in the university lab performed better on the transfer task than those tested in their preschools. A possible reason may be the greater fatigue on the final transfer task experienced by children in the preschool who were tested in between their lessons. Yet another limitation concerns the transfer task, which only required children to provide suggestions for others' behaviors rather than assessing their own application of the moral lesson. These limitations could be explored further in further research.

An interpretive concern is whether children in the pretesting condition were more likely to provide a lesson-based response in the open response task because they simply repeated answers to the pretest questions for which they received positive feedback. Although a supplementary analysis did not find significant correlations between pretest performance and open response scores or improved pretest performance across scenarios (see online Supplemental Materials), it should be noted that this possible pattern would still indicate that they had engaged in some form of abstraction by identifying similarities among the pretest scenarios and the story itself, and then provided an open-ended response free of superficial elements from either of the cases. In doing so, they might have engaged in comparative reasoning—that is, identifying similarities across cases. Also, being able to process positive feedback and then provide the desired answer could still be considered an important component of learning that children in the pretest condition benefitted from (Fyfe et al., 2023). Nonetheless, this alternative explanation may be the impetus for studies that make modifications to the existing intervention, such

as administering pretest questions with lower chances of successful responding or by assessing the effects of feedback on children's performance.

Finally, the pretesting procedure we used had limitations. Participants frequently answered the pretest questions correctly, scoring at ceiling on the "direct" pretest questions and with mean scores of 43% on the "reasoning" pretest questions (which, in the present study, refers to providing responses consistent with the intended moral lesson). Given that curiosity-based accounts of pretesting effects are predicated on low pretest question performance (Pan & Rivers, 2023), that relatively high success rate is a concern. It may, for example, have reduced children's curiosity to learn from the moral stories that followed. That possibility could also be interrogated in follow-up research.

Future Research Directions

Beyond directly addressing the aforementioned limitations, future research could address yet other areas. For instance, one important direction would be to systematically examine factors that influence the difficulty of extracting story morals (Afflerbach, 1990; Fisch et al., 2001; Narvaez, 2002). Such studies could attempt to identify these factors when testing the generalizability of the present findings to different stories. Additionally, future studies might also extend the findings to other cultural, educational, and linguistic contexts that may vary in their emphasis on moral education, values, and other respects. For example, in cultures or schools where independence is favored, young children may have less prior exposure to the intended moral lesson of working together. Pretesting effects for moral learning, if any, may as such be less likely to occur through the mechanism of activating existing knowledge schemas. A related question involves whether the present findings are generalizable beyond moral learning and, for example, to children's learning more generally. We suspect that pretesting may also be able to enhance learning in a similar fashion, but that remains to be conclusively determined.

Future studies could also consider using behavioral measures to examine if pretesting facilitates children's application of the moral lesson to their own lives. For example, after hearing the "Tall People" story, participants could be given a seemingly unrelated task where they could select playmates from a list of children and decide whether to include someone who is different from themselves and the other children (e.g., Mares & Acosta, 2010). Other paradigms for different moral themes could be adapted as well, such as administering a resource allocation task following a story about sharing (e.g., Larsen et al., 2018).

This study focused on pretesting and 5- to 6-year-olds' moral learning. As they prepare to enter formal education, future research should explore whether pretesting can still serve as a learning strategy to support older children's moral development. Given that children struggle with theme extraction until around 10 years old (e.g., Mares & Acosta, 2008) and pretesting benefits school-aged children's learning of expository texts (de Lima & Jaeger, 2020), it may also aid in moral learning at that stage. However, as children gain social experience, it might be challenging to identify suitable materials that ensure the moral stories remain novel and neutral to the majority of participants.

It might also be useful to investigate the relative efficacy of pretesting when compared to other strategies aimed at facilitating children's learning from moral stories. In this study, the fact that the control group exhibited relatively high performance on the transfer task is notable given that the control group activity—free drawing—was entirely irrelevant to learning from moral stories and, as such, arguably should have left that group at a disadvantage. In future work, comparisons could be made against questions asked during or after the story (Carnine et al., 1982; Walker & Lombrozo, 2017; for a taxonomy of practice question strategies, see Pan et al., 2024) or even against direct instruction methods where an adult explicitly states the lesson and provides children with explanations for moral actions during the story (Mares & Acosta, 2010; Walker & Lombrozo, 2017). In the pretesting literature with adults, there is emerging evidence that pretesting can be more effective than such competitive control activities as studying learning objectives (Sana et al., 2020) and even retrieval practice (Pan & Sana, 2021), but the various activities with which pretesting has been compared against remain limited. For example, if children recalled moral story details from memory before completing the dependent measures, their performance might resemble the pretesting condition, but that possibility remains untested. Future work could also reexamine the extent to which pretesting might enhance focus on an intended lesson from a moral story versus increasing overall engagement, as well as the extent to which pretesting might enhance recall of more surface-level features versus deeper understanding of a moral story.

Further, although the present study utilized moral stories with human characters engaging in realistic activities, children are also often told popular stories involving anthropomorphized animal characters or fantasy settings. These elements have been proposed to create even more difficulties for children to extract and apply the lesson to real-life scenarios due to the lower identification with the characters and their actions (e.g., Ding et al., 2023; Larsen et al., 2018). Hence, future research could investigate if pretesting improves learning from moral stories under these more difficult circumstances.

Finally, attempting to disentangle the contributions of the various components of the pretesting procedure that was used, as well as exploring yet other implementations of pretesting, may be fruitful endeavors. One possibility is the use of pretesting combined with repeated testing within the story (such as asking children to predict how a character would feel before revealing the next story scene) as in interpolated pretesting (Pan & Carpenter, 2023). There is also some evidence that questions that are higher in cognitive level can induce more thorough learning of concepts from scientific texts (e.g., St. Hilaire et al., 2019); perhaps comparing different pretest question types may be useful in the context of learning from moral stories. That approach and other variations of pretesting interventions should be considered, especially since the present method did not yield the intended transfer effects.

Significance and Implications

Despite its limitations, the present study has various theoretical and practical implications. As moral stories differ from other learning materials frequently used in past pretesting research (e.g., science texts), this study is among the first to implement the technique on more abstract forms of learning where the prescriptive message (i.e., the moral) cannot be explicitly observed in the material (Lehr, 1988; Lukens, 1982). This study also bridged the gaps in the pretesting literature by demonstrating that this technique, which has mostly been

studied in adult and student populations, also benefits preschool children. This finding thus sets the stage for further studies to investigate the effects of pretesting on children's learning in other domains.

Additionally, the finding that pretesting improves children's ability to extract the theme from moral stories presents important pedagogical implications, especially since few investigated strategies to date have shown to be effective in doing so. Instead of simply reading a story to children and expecting them to spontaneously understand its lesson, parents and educators can pose pretest questions related to the moral and encourage children to actively generate guesses to these questions before hearing the story (e.g., having to explain why it is not nice to exclude others). This novel and efficient method may also be used in place of existing popular approaches to storybook reading, such as repeatedly exposing children to the same stories (Schickedanz, 1978), which has been shown to support other learning outcomes but is less effective at teaching the moral message (e.g., Mares, 2006).

Conclusions

In sum, the present study demonstrates that pretesting improves young children's ability to extract the intended theme from moral stories. This conclusion is substantiated by the finding that pretesting significantly facilitated children's performance on the more difficult task of abstracting and generating the intended theme in their own words, even though it did not enhance their ability to choose the story's theme from a list of options. The results also revealed that pretesting did not improve children's generalization of the story's lesson to novel real-life contexts, but this nonsignificant effect could be due to methodological limitations that should be addressed in future research. Overall, the finding that pretesting improves children's extraction of the intended theme suggests that implementation of this technique could enhance the usefulness of stories as a moral education tool, helping children to learn values deemed as important for shaping their character.

References

Afflerbach, P. P. (1990). The influence of prior knowledge on expert readers' main idea construction strategies. *Reading Research Quarterly*, *25*(1), 31–46. https://doi.org/10.2307/747986

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), 3–25. https://doi.org/10.1037/amp0000191

Bennett, W. J. (1993). *The book of virtues*. Simon and Schuster.

Berner, V. D., Seitz-Stein, K., Segerer, R., Oesterlen, E., & Niklas, F. (2022). 'Good' or 'well calculated'? Effects of feedback on performance and self-concept of 5- to 7-year-old children in math. *Educational Psychology*, *42*(3), 296–315. https://doi.org/10.1080/01443410.2021.2001790

Bialystok, E. (2010). Global-local and trail-making tasks by monolingual and bilingual children: Beyond inhibition. *Developmental Psychology*, *46*(1), 93–105. https://doi.org/10.1037/a0015466

Brandão, A. C. P. (1994). *Production and comprehension of stories by children* [Unpublished master's thesis]. Federal University of Pernambuco.

Brandão, A. C. P., & Oakhill, J. (2005). "How do you know this answer?"—Children's use of text data and general knowledge in story comprehension. *Reading and Writing*, *18*(7–9), 687–713. https://doi.org/10.1007/s11145-005-5600-x

Breitwieser, J., & Brod, G. (2021). Cognitive prerequisites for generative learning: Why some learning strategies are more effective than others. *Child Development*, *92*(1), 258–272. https://doi.org/10.1111/cdev.13393

Brod, G. (2021). Generative learning: Which strategies for what age? *Educational Psychology Review*, *33*(4), 1295–1318. https://doi.org/10.1007/s10648-020-09571-9

Brown, A. L., Kane, M. J., & Echols, C. H. (1986). Young children's mental models determine analogical transfer across problems with a common goal structure. *Cognitive Development*, *1*(2), 103–121. https://doi.org/10.1016/S0885-2014(86)80014-4

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development: An International Journal of Research and Practice*, *11*(2), 73–92. https://doi.org/10.1002/icd.298

Carneiro, P., Lapa, A., & Finn, B. (2018). The effect of unsuccessful retrieval on children's subsequent learning. *Journal of Experimental Child Psychology*, *166*, 400–420. https://doi.org/10.1016/j.jecp.2017.09.010

Carnine, D., Stevens, C., Clements, J., & Kameenui, E. J. (1982). Effects of facilitative questions and practice on intermediate students' understanding of character motives. *Journal of Reading Behavior*, *14*(2), 179–190. https://doi.org/10.1080/10862968209547445

Carpenter, S. K., King-Shepard, Q., & Nokes-Malach, T. (2023). The prequestion effect: Why it is useful to ask students questions before they learn. In C. E. Overson, C. M. Hakala, L. L. Kordonowy, & V. A. Benassi (Eds.), *In their own words: What scholars and teachers want you to know about why and how to apply the science of learning in your academic setting* (pp. 74–82). Society for the Teaching of Psychology.

Carpenter, S. K., & Toftness, A. R. (2017). The effect of prequestions on learning from video presentations. *Journal of Applied Research in Memory and Cognition*, *6*(1), 104–109. https://doi.org/10.1016/j.jarmac.2016.07.014

Davis, H. L., & Pratt, C. (1995). The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology*, *47*(1), 25–31. https://doi.org/10.1080/00049539508258765

de Lima, N. K., & Jaeger, A. (2020). The effects of prequestions versus postquestions on memory retention in children. *Journal of Applied Research in Memory and Cognition*, *9*(4), 555–563. https://doi.org/10.1016/j.jarmac.2020.08.005

Ding, X. P., Cheng, J. K. T., Cheng, Q., & Heyman, G. D. (2024). An assessment of when moral stories promote children's honesty. *Applied Developmental Science*, *28*(3), 292–301. https://doi.org/10.1080/10888691.2023.2195182

Ding, X. P., Tay, C., Chua, Y. J., & Cheng, J. K. T. (2023). Can classic moral stories with anthropomorphized animal characters promote children's honesty? *Journal of Applied Developmental Psychology*, *85*, Article 101498. https://doi.org/10.1016/j.appdev.2022.101498

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

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). Pearson.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fisch, S. M., Kirkorian, H. L., & Anderson, D. R. (2005). Transfer of learning in informal education: The case of television. In J. P. Mestre (Ed.), *Transfer of Learning from a Modern Multidisciplinary Perspective* (pp. 371–393). Information Age Publishing.

Fisch, S. M., McCann Brown, S. K., & Cohen, D. I. (2001). Young children's comprehension of educational television: The role of visual information

and intonation. *Media Psychology*, *3*(4), 365–378. https://doi.org/10.1207/S1532785XMEP0304_04

Flack, Z. M., & Horst, J. S. (2018). Two sides to every story: Children learn words better from one storybook page at a time. *Infant and Child Development*, *27*(1), Article e2047. https://doi.org/10.1002/icd.2047

Foo, J. H., Yu, L., Pan, S. C., & Ding, X. P. (2025, April 29). *Effects of pretesting on children's learning of moral stories*. https://osf.io/dsc6r/

Fyfe, E. R., Borriello, G., & Merrick, M. (2023). A developmental perspective on feedback: How corrective feedback influences children's literacy, mathematics, and problem solving. *Educational Psychologist*, *58*(3), 130–145. https://doi.org/10.1080/00461520.2022.2108426

Geller, J., Carpenter, S. K., Lamm, M. H., Rahman, S., Armstrong, P. I., & Coffman, C. R. (2017). Prequestions do not enhance the benefits of retrieval in a STEM classroom. *Cognitive Research: Principles and Implications*, *2*(1), Article 42. https://doi.org/10.1186/s41235-017-0078-z

Goldman, S. R., Reyes, M., & Varnhagen, C. K. (1984). Understanding fables in first and second languages. *NABE Journal*, *8*(2), 35–66. https://doi.org/10.1080/08855072.1984.10668465

Jose, P. E., D'Anna, C. A., & Krieg, D. B. (2005). Development of the comprehension and appreciation of fables. *Genetic, Social, and General Psychology Monographs*, *131*(1), 5–37. https://doi.org/10.3200/MONO.131.1.5-37

Kochanska, G. (2002). Committed compliance, moral self, and internalization: A mediational model. *Developmental Psychology*, *38*(3), 339–351. https://doi.org/10.1037/0012-1649.38.3.339

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 989–998. https://doi.org/10.1037/a0015729

Larsen, N. E., Lee, K., & Ganea, P. A. (2018). Do storybooks with anthropomorphized animal characters promote prosocial behaviors in young children? *Developmental Science*, *21*(3), Article e12590. https://doi.org/10.1111/desc.12590

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

Lehr, S. (1988). The child's developing sense of theme as a response to literature. *Reading Research Quarterly*, *23*(3), 337–357. https://doi.org/10.2307/748046

Lehr, S. (1990). Literature and the construction of meaning: The preschool child's developing sense of theme. *Journal of Research in Childhood Education*, *5*(1), 37–46. https://doi.org/10.1080/02568549009594801

Lever, R., & Sénéchal, M. (2011). Discussing stories: On how a dialogic reading intervention improves kindergartners' oral narrative construction. *Journal of Experimental Child Psychology*, *108*(1), 1–24. https://doi.org/10.1016/j.jecp.2010.07.002

Lukens, R. (1982). *A critical handbook of children's literature* (2nd ed.). Scott Foresman.

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

Mares, M.-L. (2006). Repetition increases children's comprehension of television content—Up to a point. *Communication Monographs*, *73*(2), 216–241. https://doi.org/10.1080/03637750600693464

Mares, M.-L., & Acosta, E. E. (2008). Be kind to three-legged dogs: Children's literal interpretations of TV's moral lessons. *Media Psychology*, *11*(3), 377–399. https://doi.org/10.1080/15213260802204355

Mares, M.-L., & Acosta, E. E. (2010). Teaching inclusiveness via TV narratives in the US: Young viewers need help with the message. *Journal of Children and Media*, *4*(3), 231–247. https://doi.org/10.1080/17482798.2010.486127

McCrudden, M. T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review*, *19*(2), 113–139. https://doi.org/10.1007/s10648-006-9010-7

McDaniel, M. A., Agarwal, P. K., Huelser, B. J., McDermott, K. B., & Roediger, H. L., III. (2011). Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology*, *103*(2), 399–414. https://doi.org/10.1037/a0021782

Memory, D. M. (1983). Main idea prequestions as adjunct aids with good and low-average middle grade readers. *Journal of Reading Behavior*, *15*(2), 37–48. https://doi.org/10.1080/10862968309547482

Murphy, D. H., Little, J. L., & Bjork, E. L. (2023). The value of using tests in education as tools for learning—Not just for assessment. *Educational Psychology Review*, *35*(3), Article 89. https://doi.org/10.1007/s10648-023-09808-3

Narvaez, D. (2002). Does reading moral stories build character? *Educational Psychology Review*, *14*(2), 155–171. https://doi.org/10.1023/A:1014674621501

Narvaez, D., Bentley, J., Gleason, T., & Samuels, J. (1998). Moral theme comprehension in third graders, fifth graders, and college students. *Reading Psychology*, *19*(2), 217–241. https://doi.org/10.1080/0270271980190203

Narvaez, D., Gleason, T., Mitchell, C., & Bentley, J. (1999). Moral theme comprehension in young children. *Journal of Research in Childhood Education*, *14*(2), 145–158. https://doi.org/10.1037/0022-0663.91.3.477

National Longitudinal Survey of Youth, NLSY79 Child Surveys. (1974). *Memory for digit span*. National Longitudinal Surveys.

Pan, S. C., & Carpenter, S. K. (2023). Prequestioning and pretesting effects: A review of empirical research, theoretical perspectives, and implications for educational practice. *Educational Psychology Review*, *35*(4), Article 97. https://doi.org/10.1007/s10648-023-09814-5

Pan, S. C., Dunlosky, J., Xu, K. M., & Ouwehand, K. (2024). Emerging and future directions in test-enhanced learning research. *Educational Psychology Review*, *36*(1), Article 20. https://doi.org/10.1007/s10648-024-09857-2

Pan, S. C., & Rivers, M. L. (2023). Metacognitive awareness of the pretesting effect improves with self-regulation support. *Memory & Cognition*, *51*(6), 1461–1480. https://doi.org/10.3758/s13421-022-01392-1

Pan, S. C., & Sana, F. (2021). Pretesting versus posttesting: Comparing the pedagogical benefits of errorful generation and retrieval practice. *Journal of Experimental Psychology: Applied*, *27*(2), 237–257. https://doi.org/10.1037/xap0000345

Pan, S. C., Sana, F., Schmitt, A. G., & Bjork, E. L. (2020). Pretesting reduces mind wandering and enhances learning during online lectures. *Journal of Applied Research in Memory and Cognition*, *9*(4), 542–554. https://doi.org/10.1016/j.jarmac.2020.07.004

Pelletier, J., & Beatty, R. (2015). Children's understanding of Aesop's fables: Relations to reading comprehension and theory of mind. *Frontiers in Psychology*, *6*, Article 1448. https://doi.org/10.3389/fpsyg.2015.01448

Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General*, *143*(2), 644–667. https://doi.org/10.1037/a0033194

Pressley, M., Tanenbaum, R., McDaniel, M. A., & Wood, E. (1990). What happens when university students try to answer prequestions that accompany textbook material? *Contemporary Educational Psychology*, *15*(1), 27–35. https://doi.org/10.1016/0361-476X(90)90003-J

Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist*, *27*(1), 91–109. https://doi.org/10.1207/s15326985ep2701_7

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied*, *15*(3), 243–257. https://doi.org/10.1037/a0016496

Rickards, J. P. (1976). Interaction of position and conceptual level of adjunct questions on immediate and delayed retention of text. *Journal of Educational Psychology*, *68*(2), 210–217. https://doi.org/10.1037/0022-0663.68.2.210

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*, Article 105392. https://doi.org/10.1016/j.jecp.2022.105392

Sana, F., Forrin, N. D., Sharma, M., Dubljevic, T., Ho, P., Jalil, E., & Kim, J. A. (2020). Optimizing the efficacy of learning objectives through pretests. *CBE Life Sciences Education*, *19*(3), Article ar43. https://doi.org/10.1187/cbe.19-11-0257

Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research*, *37*(3), 409–425. https://doi.org/10.1086/651235

Schickedanz, J. A. (1978). "Please read that story again!": Exploring relationships between story reading and learning to read. *Young Children*, *33*(5), 48–55.

Schoppmann, J., Severin, F., Schneider, S., & Seehagen, S. (2023). The effect of picture book reading on young children's use of an emotion regulation strategy. *PLOS ONE*, *18*(8), Article e0289403. https://doi.org/10.1371/journal.pone.0289403

Shanahan, T. (1983). *Prequestions and predictions and children's reading comprehension* [Paper presentation]. Annual Meeting of the American Educational Research Association, Montreal, Canada.

St. Hilaire, K. J., & Carpenter, S. K. (2020). Prequestions enhance learning, but only when they are remembered. *Journal of Experimental Psychology: Applied*, *26*(4), 705–716. https://doi.org/10.1037/xap0000296

St. Hilaire, K. J., Carpenter, S. K., & Jennings, J. M. (2019). Using prequestions to enhance learning from reading passages: The roles of question type and structure building ability. *Memory*, *27*(9), 1204–1213. https://doi.org/10.1080/09658211.2019.1641209

Strouse, G. A., Nyhout, A., & Ganea, P. A. (2018). The role of book features in young children's transfer of information from picture books to real-world contexts. *Frontiers in Psychology*, *9*, Article 50. https://doi.org/10.3389/fpsyg.2018.00050

Vaahtoranta, E., Suggate, S., Jachmann, C., Lenhart, J., & Lenhard, W. (2018). Can explaining less be more? Enhancing vocabulary through explicit versus elaborative storytelling. *First Language*, *38*(2), 198–217. https://doi.org/10.1177/0142723717737452

van den Broek, P., Kendeou, P., Kremer, K., Lynch, J., Butler, J., White, M. J., & Lorch, E. P. (2005). Assessment of comprehension abilities in young children. In S. Stahl & S. Paris (Eds.), *Children's reading comprehension and assessment* (pp. 107–130). Erlbaum.

van den Broek, P., Lorch, E. P., & Thurlow, R. (1996). Children's and adults' memory for television stories: The role of causal factors, story-grammar categories, and hierarchical level. *Child Development*, *67*(6), 3010–3028. https://doi.org/10.2307/1131764

van den Broek, P., Lynch, J. S., Naslund, J., Ievers-Landis, C. E., & Verduin, K. (2003). The development of comprehension of main ideas in narratives: Evidence from the selection of titles. *Journal of Educational Psychology*, *95*(4), 707–718. https://doi.org/10.1037/0022-0663.95.4.707

Walker, C. M., & Lombrozo, T. (2017). Explaining the moral of the story. *Cognition*, *167*, 266–281. https://doi.org/10.1016/j.cognition.2016.11.007

Williams, J. P., Lauer, K. D., Hall, K. M., Lord, K. M., Gugga, S. S., Bak, S.-J., Jacobs, P. R., & deCani, J. S. (2002). Teaching elementary school students to identify story themes. *Journal of Educational Psychology*, *94*(2), 235–248. https://doi.org/10.1037/0022-0663.94.2.235

Wu, S., & Tan, C. (2021). A holistic model of competence: Curriculum reforms for pre-school education in Singapore. In W. Zhao & D. Tröhler (Eds.), *Euro-Asian encounters on 21st-century competency-based curriculum reforms* (pp. 211–227). Springer.

Yu, L., Hu, J. T., Arden-Gardner, L. S., Ding, X. P., & Pan, S. C. (2025). *Prequestioning with immediate feedback promotes learning in young children* [Manuscript submitted for publication]. Department of Psychology, National University of Singapore.

Received January 16, 2025
Revision received April 26, 2025
Accepted April 28, 2025 ■