The moral, or the story?

Changing children’s distributive justice preferences through social communication

Joshua Rottman\textsuperscript{a}, Valerie Zizik\textsuperscript{a}, Kelly Minard\textsuperscript{a}, Liane Young\textsuperscript{b}, Peter R. Blake\textsuperscript{c}, & Deborah Kelemen\textsuperscript{c}

\textsuperscript{a} Department of Psychology, Franklin & Marshall College
P.O. Box 3003, Lancaster, PA 17604

\textsuperscript{b} Department of Psychology, Boston College
140 Commonwealth Avenue, Chestnut Hill, MA 02467

\textsuperscript{c} Department of Psychological and Brain Sciences, Boston University
64 Cummington St., Boston, MA 02215

**Corresponding Author**

Joshua Rottman
Department of Psychology
Franklin & Marshall College
P.O. Box 3003, Lancaster, PA 17604
E-mail: jrottman@fandm.edu
Phone: 717-358-4874
Abstract
Can social communication alter children’s preexisting inclinations toward equality-based or merit-based forms of resource distribution? Six- to eight-year-old children’s ($N = 248$) fairness preferences were evaluated with third-party distribution tasks before and after an intervention. Study 1 indicated that stories about beavers dividing wood had no impact on children’s fairness preferences, while Study 2 indicated that brief, direct testimony was robustly influential. Study 3 matched storybooks and testimony in content, with each discussing a situation resembling the distribution task, and both formats exerted a significant impact on children’s fairness preferences that persisted across several weeks. Interventions preaching the superiority of equality-based fairness were particularly effective. Overall, children’s distributive justice preferences can be profoundly influenced by easily transferable forms of social communication.

Keywords: storybooks; testimony; fairness; distributive justice; moral development
The Moral, or the Story?

Changing Children’s Distributive Justice Preferences Through Social Communication

1. Introduction

The witch in Sondheim’s Into the Woods warns: “Careful the things you say; children will listen.” This aptly reflects the popular belief that children readily absorb anything they hear, whether they are told directly or through media (e.g., television, movies, or books) – particularly when learning moral values. Certainly, children are often credulous, believing what adults tell them even if it conflicts with their firsthand experience (see Jaswal, 2013). However, children’s sponge-like credulity is primarily bound to situations of ambiguity for which they have no prior knowledge. Children are considerably more discerning and prone to reject adults’ claims when they already hold a countervailing belief or when they have other reasons not to trust an informant (see Harris, 2012; Koenig & Sabbagh, 2013; Mills, 2013; Sobel & Kushnir, 2013). An open question therefore remains: how receptive are children to social communication that aims to influence their preexisting moral commitments?

In the present research, we investigated whether storybooks and direct testimony – two common sources of information about the world – could lead children to value one form of fairness (e.g., equality) above another (e.g., merit), as measured by third-party resource distributions. This contrast between equality-based fairness and merit-based fairness has been prevalent throughout history and prominently fuels contemporary political debates and ideologies in the United States (Dworkin, 2000). It is therefore important to address the extent to which children adopt a reliable stance on how to justly allocate resources, and whether certain social influences can disrupt this stance.
1.1. The development of distributive justice

Since Piaget (1932), developmental psychologists have demonstrated that children possess at least two distinct ideas of how resources should be properly apportioned: according to equality (with everybody getting the same amount) or according to merit (with more deserving recipients earning more rewards). Early research on the development of fairness beliefs was focused on the influences of cognitive maturation and logical or mathematical reasoning, positing that children pass through stable cognitive stages that lead to changes in fairness concepts (for reviews, see Damon, 1977; Hook & Cook, 1979). This body of literature yielded evidence that equality-based fairness is robust during early and middle childhood and is then replaced by merit-based fairness, which seems to naturally emerge during late childhood or early adolescence (Almås, Cappelen, Sørensen, & Tungodden, 2010; Damon, 1975; Piaget, 1932).

However, more recent research has yielded suggestive evidence for a less stage-like view of the development of fairness beliefs (e.g., Blake et al., 2015). These findings have supported theories that children’s moral beliefs are shaped at least in part by learning from discourse transmitted by moral authorities within a given cultural milieu (e.g., Shweder, Mahapatra, & Miller, 1987; Snarey, 1985). Some studies have found that children’s distribution patterns are malleable, changing with context in addition to age (Huntsman, 1984; McGillicuddy-de Lisi, Watkins, & Vinchur, 1994; Sigelman & Waitzman, 1991). Additionally, recent research has demonstrated that the concepts of equality and merit are simultaneously available even to very young children. In looking time paradigms, infants are sensitive both to principles of equality (Geraci & Surian, 2011; Ziv & Sommerville, 2017) and to principles of merit (Sloane, Baillargeon, & Premack, 2012). Even three- and four-year-old children have been found to apply principles of merit when distributing resources (Baumard, Mascaro, & Chevallier, 2012;
Kanngiesser & Warneken, 2012; Nelson & Dweck, 1977), although tendencies toward merit-based distributions may be especially characteristic of Westernized, industrialized cultural contexts (Jara-Ettinger, Gibson, Kidd, & Piantadosi, 2016; Schäfer, Haun, & Tomasello, 2015).

Together, the early emergence of both equality- and merit-based fairness and the context-dependence of distributive justice suggest that favoring one form of fairness over another requires a sociocultural, rather than purely cognitive, explanation. Given that children can appreciate the principles of equality and merit from an early age, cognitive abilities cannot fully account for individual differences in children’s tendencies to rely on one or the other form of resource distribution. Despite extensive cognitive developmental research on children’s third-party resource distribution tendencies, research is lacking on the forms of social influence that shape these preferences. Could storybooks or testimony bring about shifts in preexisting fairness preferences, leading children to revise their tendencies either to provide people with equal amounts of resources or to provide more industrious people with more resources? The present research investigates whether patterns of resource distribution are amenable to being changed through such forms of social communication. This research also tested whether social communication is more effective for instilling preferences for equality, which is likely an initial default stance (Baumard et al., 2012; Sigelman & Waitzman, 1991; Ziv & Sommerville, 2017), or for merit, given that children typically progress from preferring equality to preferring merit rather than the opposite (Hook & Cook, 1979).

1.2. Learning moral lessons from storybooks and other forms of social communication

Throughout history, narratives have been utilized to shape moral beliefs. This is evidenced in religious texts such as the Bible and the Ramayana, as well as in secular stories such as Aesop’s fables. Anecdotally, stories have had tremendous effects in ameliorating
prejudice and discrimination, as when enslaved African-Americans were treated with greater moral concern after the wide circulation of *Uncle Tom’s Cabin* (Bloom, 2010). Accounts such as these have led several researchers to identify narrative as a primary cultural instrument for shaping and perhaps revising moral intuitions (Bloom, 2010; Haidt & Joseph, 2007; Tappan & Brown, 1989). Indeed, research has shown that stories can lead to powerful changes in adults’ moral attitudes (Green & Brock, 2000; Mazzocco, Green, Sasota, & Jones, 2010). The hypothesis that storybooks should be a viable form of social communication for influencing children’s beliefs about distributive justice is also suggested by research indicating that illustrated storybooks are frequently effective vehicles of belief change in childhood (e.g., Kelemen, Emmons, Schillaci, & Ganea, 2014; for reviews, see Hopkins & Weisberg, 2017; Strouse, Nyhout, & Ganea, 2018).

Two studies have investigated the effectiveness of storybooks in producing changes in children’s resource allocations. In one study, 6- to 9-year-old children became more generous in a first-person altruism task (the Dictator Game) after being read a storybook about a child who exerted self-control (Steinbeis & Over, 2017). In another study, 4- to 6-year-old children became more generous in the Dictator Game when they were read a story about a child who learned to share, although not when they were read an otherwise identical story about an anthropomorphic raccoon who learned to share (Larsen, Lee, & Ganea, 2017). Both studies therefore demonstrated that storybook interventions can affect children’s altruistic giving. However, neither study attempted to change children’s abstract fairness preferences; rather than targeting belief revision, participants were led to act in accordance with their preexisting beliefs that generosity was good. Influencing beliefs about fair resource distribution requires an approach that targets specific fairness principles, ideally when children do not stand to benefit themselves.
Despite some successes in teaching children moral lessons through stories, as described above, many written narratives and television shows have been ineffective, often as a result of poor comprehension of the intended message or a failure to transfer the information to a different context (Mares & Acosta, 2008, 2010; Narvaez, Gleason, Mitchell, & Bentley, 1999). For example, upon watching a television episode about a three-legged dog (intended to teach tolerance for people with disabilities), most kindergarteners took the story at face value and did not appropriately generalize the message to humans – and in some cases, they became more intolerant of those with disabilities, thus internalizing exactly what the story aimed to overcome (Mares & Acosta, 2008). Additionally, a recent set of attempts to teach children honesty through popular storybooks such as Pinocchio found that only one of four books was effective (Lee et al., 2014), and it is unclear whether this single success resulted from the message contained in the storybook itself or from the experimenter’s prompting of children to act like the protagonist.

Overall, the conflicting literature on the efficacy of moral learning from storybooks suggests that children may derive greater benefits from more direct and straightforward forms of verbal communication (henceforth, simplified to “testimony”). Unlike stories, which are largely intended to entertain, the primary purpose of testimony is to convey information. A number of studies have found that children’s moral beliefs and prosocial behaviors can be influenced by various forms of adult testimony (McAuliffe, Raihani, & Dunham, 2017; Rosenhan, Frederick, & Burrowes, 1968; Rottman, Young, & Kelemen, 2017; Rushton, 1975; Sagotsky, Wood-Schneider, & Konop, 1981; also see Eisenberg, Fabes, & Spinrad, 2006; Harris, 2012). Therefore, in addition to examining the impact of storybooks on fairness preferences (Studies 1 and 3), we investigated whether a brief direct statement (Study 2) or more extended testimony (Study 3) would affect children’s fairness preferences.
1.3. Overview of the present research

Our research investigated whether storybooks and testimony are effective tools for influencing children to revise their existing preferences for either equality-based or merit-based forms of resource distribution. First, to estimate the stability of children’s fairness preferences, we conducted a baseline study to test children’s tendencies to change their patterns of resource distribution after an irrelevant intervention. Then, in our primary studies, we examined children’s tendencies to change their patterns of resource distribution after being exposed to an intervention that was designed to convince them of the superiority of their disfavored form of fairness.

In each of the three primary studies, participants were assigned to one of four conditions. In Studies 1 and 2, these resulted from a 2 (Fairness Type: Merit vs. Equality) X 2 (Appeal: Emotional vs. Reasoned) design. In Study 3, these resulted from a 2 (Fairness Type: Merit vs. Equality) X 2 (Intervention: Storybook vs. Testimony) design. We tested participants both before and after an intervention to determine whether social communication could alter the fairness preferences children already possessed. To measure fairness preferences, a third-person distribution task was used. This not only allowed us to investigate the socialization of moral principles, but also eliminated the potential that selfishness could influence the results. When

---

1 Because a distinction between emotional and reasoned appeals has been frequently emphasized in the literatures on persuasion (Petty & Cacioppo, 1986), moral psychology (Haidt, 2001), and narrative transportation (Mazzocco et al., 2010), we tested whether emotional appeals are more effective than reasoned appeals in Studies 1 and 2. As this was a secondary interest, and as there were no effects of this variable, we have eliminated discussion of the Appeal manipulation in the main text. All relevant details can be found in Supplementary Materials.
children stand to gain or lose from their distributions of resources, they tend to be strategically self-interested (Fehr, Bernhard, & Rockenbach, 2008; Larsen et al., 2017; Shaw et al., 2014; Sheskin, Bloom, & Wynn, 2014; Smith, Blake, & Harris, 2013; Steinbeis & Over, 2017) and do not always behave in accordance with their principles (Blake, McAuliffe, & Warneken, 2014). These considerations are not relevant to impartial third-person allocations, which may more directly reflect abstract beliefs about justice.

We tested six- and seven-year-old children in Studies 1 and 2 because, during this period of development, children begin gravitating away from heavily weighting equality-based forms of distribution and move toward merit-based forms of distribution (e.g., Damon, 1975; Hook & Cook, 1979; Leventhal, Popp, & Sawyer, 1973). Additionally, findings that children are overly literal in their interpretation of moral stories have been primarily constrained to preschool-aged participants (see Narvaez, 2002), and older children have enhanced abilities to extract abstract lessons from storybooks (Emmons, Lees, & Kelemen, 2018). Due to a preponderance of equality distributors in the pre-test phase in Studies 1 and 2, Study 3 broadened the age range to include eight-year-olds. In all studies, we report all measures, conditions, and data exclusions, either in the main text or in the Supplementary Materials. All sample sizes were determined in advance.

2. Baseline Study

Because the stability of children’s fairness preferences is largely unknown, we conducted an additional study to obtain a point estimate of “chance” responding. This control study involved an irrelevant intervention, which allowed us to estimate the baseline rate at which children can be expected either to remain steadfast in their patterns of resource distribution between a pre-test and a post-test, or to randomly oscillate between different kinds of resource
distributions (i.e., equality-based fairness or merit-based fairness) across these two similar
distribution tasks.

2.1. Method

2.1.1. Participants

Participants were 32 six- to eight-year-olds (14 female; \(M_{\text{age}} = 7.41; SD_{\text{age}} = 0.93\)) who were recruited from a small city in the northeastern United States via a participant database and tested in a college laboratory \((N = 12)\), or who were tested in a nearby school \((N = 20)\). Participants were primarily White and middle- to high-SES. Three additional participants were tested but were replaced due to failures to pass a comprehension check. As all participants either divided stickers based on merit or based on equality in both the pre-test and the post-test, no other participants were excluded from analyses.

2.1.2. Materials and Procedure

2.1.2.1. Pre-test and Post-test. As a measure of initial fairness preferences, participants were presented with a distribution task before hearing the storybook. A second distribution task was then administered after the experimental intervention. The crucial dependent measure was whether participants changed their tendencies to distribute resources based on equality (by giving four prizes to each child) or based on merit (by giving more prizes to the child who completed more of a task).

The distribution task, adapted from Leventhal et al. (1973), took the form of a timed “work task”. Participants first completed a version of this task themselves, and they were told that time had run out after they had completed exactly 50% of the task (such that the performance of all participants was exactly the same). Because this procedure was repeated in the pre-test and the post-test, two separate work tasks were used. These were functionally similar
but differed in superficial properties. One involved adhering colored discs to a strip of colored paper, and the other involved stamping rubber stamps below corresponding animal pictures. The order of these two tasks was counterbalanced across subjects, as were the resources being distributed (either smiley-face stickers or temporary tattoos).

After participants gained experience with the work task, they were told about two other children (matched for gender and age), who had participated earlier but needed to leave suddenly and were not able to receive any prizes. It was revealed that one of these children had completed 25% of the task in the time allotted, while the other had completed 75% of the task. This discrepancy was presented visually and through counting the number of discs or stamps that had been applied to the other children’s strips of paper. The ratio was 15:5 for the discs task and 12:4 for the stamps task. At this point, participants were told that their help was needed in determining the right number of prizes to allocate to each child, and they were provided with eight resources to distribute into envelopes. A novel second experimenter administered the second task after the intervention had taken place and after playing a brief distractor game of “I Spy” with the participant in order to make this second portion of the testing session appear “totally different” from the first portion (which children were told was the case). After participants finished distributing resources on each occasion, they were asked to explain their reasoning for their particular division of stickers or tattoos.

2.1.2.1. Intervention. The baseline intervention consisted of a 357-word storybook, which was factual rather than moralistic, and which did not mention fairness, resources, or anything else related to equality or merit. Seven illustrations of beavers (also used in the Study 1 storybooks), which did not depict any form of resource distribution or indications of equality or merit, were used to create an educational storybook that provided various a range of true facts
about beavers (e.g., “Beavers have see-through eyelids that let them see things underwater” and “Beavers have long teeth that never stop growing”).

Following the second distribution task, participants were asked to recount one thing that they learned about beavers from the storybook. This allowed us to ensure attentiveness and comprehension of the book (and to analyze only data from participants who demonstrated comprehension).

2.2. Results

Across both the pre-test and the post-test, 100% of participants either divided stickers based on merit (i.e., they gave more stickers to the child who was shown to be more productive in the work task) or divided stickers equally; none gave more stickers to the child who was less productive. Children tended to be equality-distributors (62.5%) rather than merit-distributors (37.5%) at pre-test, but a binomial test indicated that this trend was non-significant, $p = .215$. A logistic regression indicated that changes in patterns of resource distribution between pre-test and post-test did not significantly differ across children who initially divided by equality or by merit, $b = 1.04, p = .234$. Critically, this study provided an estimate of the baseline rate of change in patterns of resource distribution between the pre-test and the post-test, in the absence of exposures to a relevant intervention. Overall, 7 of the 32 participants (21.88%) changed their pattern of distribution between the pre-test and the post-test.

2.3. Discussion

This Baseline Study found that more than three-quarters of participants retained the same pattern of resource distribution after an irrelevant intervention, demonstrating that children tend to have stable fairness preferences by middle childhood. This study additionally suggested a benchmark for determining the effectiveness of our interventions in Studies 1–3. Because
participants in this study were largely, but not wholly, consistent in their patterns of resource
distribution across the pre-test and post-test, it is evident that neither 0% (floor) nor 50% (pure
chance) would be appropriate points of comparison. Instead, we can reasonably expect that
approximately 21.88% of children will tend to alter their distribution patterns randomly,
regardless of whether they are exposed to a relevant intervention. We therefore use this
percentage as a point estimate for chance-level responding, to which we compare rates of change
in the studies described below.

3. Study 1

Study 1 examined whether storybooks can shift children’s fairness preferences. The
stories used in the intervention were modeled after existing popular storybooks, and as such they
featured animals as the main characters – a pervasive component of children’s storybooks
(McCrindle & Odendall, 1994).

3.1. Method

3.1.1. Participants

Participants were 70 six- and seven-year-old children (34 female; $M_{age} = 7.13$; $SD_{age} =
0.49$) from a large city in the northeastern United States who were recruited via a large
participant database and tested in a university laboratory ($N = 11$) or who were recruited and
tested in local elementary schools and summer camps ($N = 59$). Participants were primarily
White and middle- to high-SES. One additional child was tested but not included in the dataset
due to an experimenter error. Because no participants gave a greater number of resources to the
less productive child during either pre-test or post-test (i.e., they all divided prizes based on
either merit or equality), all other participants were retained for analyses.
Given that the study design involved attempting to shift children’s initial preferences for
equality or merit through a storybook intervention, participants were not randomly assigned to
the Equality or Merit interventions, but were rather assigned to a particular Fairness Type based
on the preferences that they demonstrated in their pre-test distributions (as in Bandura &
McDonald, 1963). This ensured that each participant was presented with an argument that ran
contrary to his or her initial mode of allocation (i.e., equality-distributors at pre-test were
assigned to hear a storybook that advocated for merit, while merit-distributors at pre-test were
assigned to hear a storybook that advocated for equality). Given this targeted intervention
approach, over-sampling was necessary in order to have sufficient sample sizes in each
condition. A stopping rule, decided prior to data collection, dictated that testing would cease
after at least 16 children had been assigned to each cell within the 2 (Fairness Type) X 2
(Appeal) design (as described in Footnote 1, children were randomly assigned to one of two
kinds of Appeal within each Fairness Type). In order to achieve this minimum sample size in the
condition preaching Equality through Emotional appeals, six extra participants needed to be
tested (an additional six participants divided by equality at pre-test and were therefore assigned
to the condition advocating Merit through Emotional appeals). Here, as well as in Studies 2 and
3, the same pattern of findings is found when excluding these additional children and reanalyzing
the data with a reduced sample of 64 participants (see Supplementary Materials).

3.1.2. Materials and Procedure

3.1.2.1. Pre-test and Post-test. Participants were presented with the same pre-test and
post-test as in the Baseline Study, such that they were again asked to distribute resources
between one child who had completed 25% of the work task and another child who had
completed 75% of this task. As was the case in the Baseline Study, the experimenters did not
specify a reason for this discrepancy in the children’s performance, making it unclear whether differences in ability or differences in effort were at play. This ambiguity was intentional, as it allowed the Equality intervention to impact fairness preferences by highlighting that individuals differ in their abilities to accomplish tasks, and claiming that we shouldn’t fault hard-working individuals for reduced outputs that are no fault of their own. Similarly, this allowed the Merit interventions to impact fairness preferences by highlighting that individuals differ in their tendencies to be diligent, and claiming that we shouldn’t reward lassadaisical individuals for their relatively poor outputs. Again, a second experimenter administered the post-test to reduce potential demand effects.

3.1.2.2. Intervention. The stories were custom written by the first author to maximize control of the relevant variables. The illustrated storybooks were each 357 words in length (exactly matching the length of the storybooks from the Baseline Study). These stories depicted a society of beavers who needed to decide how to distribute the wood they cut down. In each case, the beavers initially distributed wood in a way that matched the participant’s pattern of distribution. The beavers subsequently realized that it was better to instead divide their wood in accordance with a different principle – either equality or merit, whichever opposed the participant’s distribution – and they unanimously decided to enact this change. (See Fig. 1 for excerpts; the full set of materials is available upon request.) After the book was read, participants were asked to retell the story to the first experimenter in order to ensure full comprehension, and any inaccuracies were corrected as needed. This was relatively infrequent, as most participants demonstrated an understanding of the stories without the need for any corrections or reminders.
**Emotional:** “Benny watched how the beavers who didn’t work very hard still got the same amount of wood as the hard-working beavers, and because some beavers got more wood than they worked for, they all had the same sized dams. It made Benny so upset that he couldn’t sleep night after night.”

**Reasoned:** “The reason beavers should keep their own wood is that harder-working beavers should not be forced to give up some of the wood they cut down, and it wouldn’t be right for beavers that don’t work very hard to get extra wood that they didn’t cut down themselves.”

**Emotional:** “Benny watched how the beavers who had longer teeth had an easier time getting wood than the shorter toothed beavers, and because they didn’t give any of their own wood away to the beavers with short teeth, they had much bigger dams. It made Benny so upset that he couldn’t sleep night after night.”

**Reasoned:** “The reason beavers should split up their wood equally is that beavers with smaller teeth can’t cut down as much wood as other beavers, and it wouldn’t be right for them to end up with smaller dams just because they can’t get as much wood as longer-toothed beavers.”

---

**Fig. 1.** Sample excerpts and illustrations from each of the storybooks that were used in Study 1.

---

3.1.2.3. **Exploratory Questions.** After the primary measures were administered, participants were asked to rate, on a scale from 1 (really bad) to 4 (really good), the degree to which it is good for all beavers to split up their wood equally so every beaver gets the same amount. They were also separately asked to rate, on the same scale, the degree to which it is good for beavers to keep exactly the amount of wood they work to cut down. Two additional post-test measures, both intended to reduce the distance of the transfer required from the storybook, were added after data collection began, and were administered to a subset of the children at the very end of the study session. These exploratory measures are described fully in the Supplementary Materials.
3.2. Results

3.2.1. Primary Analyses

Participants were coded as having changed (1) or not changed (0) their type of fairness distributions between equality and merit from pre-test to post-test. Thus, regardless of whether they altered the exact quantity or ratio of stickers they distributed, they were only considered to have changed if their distribution pattern reflected a different normative principle. Overall, 16/70 (22.86%) of participants changed their distribution patterns after being read a storybook. This closely resembles the rate of change we found for the irrelevant intervention in our Baseline Study (21.88%), as confirmed by a one-sample binomial test using this percentage as a point estimate, $p = .885$. Tendencies for changing distribution patterns did not differ from the Baseline Study in either the Equality conditions (21.88% change) or the Merit conditions (23.68% change), as confirmed by one-sample binomial tests, $ps > .84$ (see Fig. 2).

A logistic regression indicated that children were not differentially likely to change their style of resource distribution across the two Fairness Types (Merit Intervention vs. Equality Intervention), $b = -0.10$, $p = .858$. Adding Age in Months and Gender to the model also did not yield any significant predictors of changes in resource distribution, $ps > .47$. 
**Fig. 2.** The percentage of participants who changed their patterns of resource distribution from the first distribution task (pre-test) to the second distribution task (post-test) in each Study (with both timepoints displayed for Study 3), split by Fairness Type and Intervention. The dashed line indicates “chance” performance, as suggested by the point estimate we obtained in the Baseline Study.
3.2.2. Secondary Analyses and Justifications of Resource Distribution

A majority of participants (77.14%) gave responses that were consistent with the book when asked which kind of fairness they endorsed for the beavers. Specifically, participants who heard the Equality books believed that it was better for beavers to divide their wood equally, as rated on a 1–4 Likert scale ($M = 3.69$, $SD = 0.82$), than to keep the amount of wood they worked to cut down ($M = 1.66$, $SD = 0.83$), $t(31) = 8.06$, $p < .001$. Conversely, participants who heard the Merit books believed that it was better for beavers to keep the amount of wood they worked to cut down ($M = 3.42$, $SD = 0.86$) than to divide their wood equally ($M = 2.61$, $SD = 1.03$), $t(37) = 3.09$, $p = .004$. These analyses suggest that most of the participants successfully learned the moral of the story as it applied to beavers, despite their general failure to transfer the lesson of the storybook to the post-test resource distribution task.

We analyzed participants’ justifications to determine whether children tended to change the content of their reasoning in accordance with changes in resource distributions. Each justification was assigned a single code based on which of five predefined categories seemed most representative (see Table 1). Participants’ justifications tended to accord with their patterns of resource distribution; merit distributors typically focused on the amount of output that the hypothetical children produced in the work task, while equality distributors typically focused on moral principles (see Fig. 3 and Supplementary Materials for details). The small number of participants who changed their distributions also tended to change the content of their justifications, unlike participants who did not change their distributions.
Table 1

*Coding for justifications, as applied across all three studies.*

<table>
<thead>
<tr>
<th>Justification Type</th>
<th>Representative Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Consideration of Outputs</strong></td>
<td>“This girl did more stamps and this girl did less stamps.” “They both did a really good job.”</td>
</tr>
<tr>
<td>The children should be rewarded based on the quantity that they produced.</td>
<td></td>
</tr>
<tr>
<td><strong>Consideration of Inputs</strong></td>
<td>“This girl was a little faster and neater.” “They look like they both tried really hard.”</td>
</tr>
<tr>
<td>The children should be rewarded based on how much effort or skill they devoted to the task.</td>
<td></td>
</tr>
<tr>
<td><strong>Consideration of Recipients’ Welfare</strong></td>
<td>“The other would cry if they didn’t get the same amount.” “So they don’t have to fight over it.”</td>
</tr>
<tr>
<td>The distribution would maximize happiness or minimize sadness/victimhood.</td>
<td></td>
</tr>
<tr>
<td><strong>Consideration of Moral Principles</strong></td>
<td>“If we sorted it any other way it wouldn’t be fair.” “It wouldn’t be fair if one got less than the other one.”</td>
</tr>
<tr>
<td>The distribution should be based upon an abstract deontic principle of justice or fairness.</td>
<td></td>
</tr>
<tr>
<td><strong>Other/Uncodable</strong></td>
<td>“I don’t remember.” “Because there’s eight and four plus four equals eight.”</td>
</tr>
</tbody>
</table>
Fig. 3. Proportions of Justification Types produced across studies, split by Fairness Type and Distribution Task (Pre-Test vs. Post-Test) and Change.
3.3. Discussion

In this study, storybooks about beavers dividing wood were designed to alter children’s inclinations toward favoring equality or merit. Although these stories resembled many popular children’s books that are marketed explicitly for the purpose of moral education, more than three-quarters of participants remained steadfast in their preferences for equality-based or merit-based forms of distributive justice after being read a story that advocated for a different form of resource distribution. Children’s overall rate of change was nearly identical to the rate of change found in the Baseline Study, which employed irrelevant storybooks that merely provided facts about beavers’ biological and behavioral characteristics. This indicates that the small number of participants who changed their pattern of resource distribution in accordance with the storybook’s lesson was likely due to random chance.

A plausible explanation for our null findings is that the storybooks were ineffective due to children’s failures to abstract the moral message being conveyed in the stories and their difficulties in analogically mapping the story’s lesson onto the distribution task (Mares & Acosta, 2008; Narvaez, 2002; Walker & Lombrozo, 2017). This is bolstered by findings that, even though participants’ own distributions of resources to other children tended to be at odds with the message being conveyed in the storybooks, participants tended to believe that beavers should divide wood in the way that the storybooks suggested. Participants also tended to succeed on other exploratory tasks, which involved scenarios that were more directly analogous to the content in the storybooks (e.g., involving distributions of nuts amongst squirrels or distributions of firewood amongst humans; see Supplementary Materials). Because the distance of transfer was reduced on these tasks, taken together, the pattern of findings suggests that participants’ failures to apply the storybooks’ lessons to the post-test resource distribution task was not due to
a problem of comprehension but was rather due to a problem of transfer. Thus, despite findings that even very young children are capable of abstract transfer (Brown, 1990), it may take somewhat longer for children to attend to deep structural features of stories (Daehler & Chen, 1993). In general, storybooks have been shown to promote “near transfer” much more effectively than “far transfer” in children (Emmons et al., 2018; Ganea, Canfield, Simons-Ghafari, & Chou, 2014; Richert & Smith, 2011; Walker, Gopnik, & Ganea, 2015), and as such they may only facilitate moral decision-making when analogical processing is relatively simple.

Therefore, a major challenge is faced when using storybooks to actuate moral belief change. Because far transfer is often required when applying lessons from popular stories, as these stories cannot be easily tailored to a particular situation at hand, it is likely that moralistic storybooks are inherently difficult for children to generalize to their everyday lives. This is unlike the case of naturalistic speech, which can be readily tailored to direct children’s attention to relevant features of a particular context. As such, testimony may represent an underappreciated alternative form of social communication that could more powerfully impact children’s fairness preferences.

4. Study 2

To test the effectiveness of a more direct means of cultural transmission for impacting children’s fairness preferences, Study 2 employed an intervention involving testimony. Given that children are typically less credulous in contexts that do not require them to defer to adults in order to acquire information, and in contexts where they already have stable commitments (Harris, 2012; Schillaci & Kelemen, 2014; Sobel & Kushnir, 2013), it is an open question whether testimony can motivate belief change.
4.1. Method

4.1.1. Participants

Participants were 101 six- and seven-year-old children (44 female; $M_{age} = 7.08$; $SD_{age} = 0.50$) from a large city in the northeastern United States who were recruited via a large participant database and tested in a university laboratory ($N = 34$) or who were recruited and tested in local elementary schools and summer camps ($N = 67$). Participants were primarily White and middle- to high-SES. An additional 9 participants gave a greater number of stickers to the less productive child during either the pre-test or the post-test. These children were excluded from all analyses, as they could not be readily classified as Merit or Equality distributors and thus there was not a clearly appropriate intervention condition to which they could be assigned.

As in Study 1, participants were assigned to the Equality or Merit conditions based on their pre-test distributions, thus ensuring that each child was presented with an argument that ran contrary to their initial mode of allocation. Despite being recruited from the same population, participants in Study 2 were more prone to equal distributions at pre-test than participants in Study 1, which meant that 37 additional children needed to be tested in the Merit conditions before 16 participants were obtained in each cell of the 2 (Fairness Type) X 2 (Appeal) design.

4.1.2. Materials and Procedure

The pre-test and post-test were identical to those used in Study 1. All participants were presented with a pre-test distribution task to establish an initial measure of fairness preferences. The first experimenter then provided brief testimony preaching the opposing form of distributive justice, after which a novel second experimenter entered the room and administered a distractor task and then a post-test distribution task. Again, the crucial dependent measure was whether participants changed their preference to distribute the resources based on equality or merit.
4.1.2.1. Intervention. The first experimenter provided testimony consisting of short declarations (59 words) presented in conversational language, without illustrations or written words. These appealed to recipients’ feelings (e.g., “dividing up stickers [in the way you demonstrated] makes these girls feel really upset”) or abstract moral principles (e.g., “each girl should have as many stickers as she deserves based on what she did to earn them”); again, as the type of Appeal (Emotional vs. Reasoned) did not impact the results, it is not discussed further here. The first experimenter left the room immediately after providing the testimony, and did not reappear again during the study session, thereby reducing potential demand effects.

Just as often happens in real life when adults directly convey moral lessons to children, the testimony focused on the situation at hand. The content of this testimony involved language that was adapted directly from the storybooks in Study 1, but which referred to human children in the immediate context. In particular, the testimony distilled the main lessons from the Study 1 stimuli without adding any additional content; a detailed allegory about beavers dividing wood was replaced with stark statements indicating either that it is much better for harder workers to get more resources or that it is much better for resources to be divided equally. (See Table 2 for the full set of materials.)
Table 2

The full, verbatim testimony that was used in Study 2.

<table>
<thead>
<tr>
<th>MERIT INTERVENTION</th>
<th>EQUALITY INTERVENTION</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Emotional:</strong> “Another way to divide up stickers is for the harder-working boy to get more stickers than the less hard-working boy. That’s a much better way of dividing up stickers, because it would have made the boys much happier if they got exactly the amount they worked for. Dividing up stickers any other way makes these boys feel really upset.”</td>
<td><strong>Emotional:</strong> “Another way to divide up stickers is for each boy to get an equal number of stickers. That’s a much better way of dividing up stickers, because it would have made the boys much happier if each one got exactly the same amount as the other boy. Dividing up stickers any other way makes these boys feel really upset.”</td>
</tr>
<tr>
<td><strong>Reasoned:</strong> “Another way to divide up stickers is for the harder-working boy to get more stickers than the less hard-working boy. That’s a much better way of dividing up stickers, because each boy should have as many stickers as he deserves based on what he did to earn them. Dividing up stickers any other way would not make any sense.”</td>
<td><strong>Reasoned:</strong> “Another way to divide up stickers is for each boy to get an equal number of stickers. That’s a much better way of dividing up stickers, because each boy should get the same amount of stickers whether or not he’s able to quickly put down lots of discs. Dividing up stickers any other way would not make any sense.”</td>
</tr>
</tbody>
</table>

4.2. Results

4.2.1. Primary Analyses

Participants were coded as having changed (1) or not changed (0) their type of fairness distributions between equality and merit from pre-test to post-test. Collapsing across conditions, a one-sample binomial test (comparing to a baseline rate of 21.88% change) found that testimony reliably led to changes in children’s distribution patterns from pre-test to post-test, \( p < .001 \), with 67.33% of participants changing their distribution patterns after testimony was presented. Children were more easily swayed into endorsing equality (81.25% change from merit) than endorsing merit (60.87% change from equality), although participants were well above Baseline levels in both conditions, as confirmed by one-sample binomial tests, \( ps < .001 \) (see Fig. 2).

A logistic regression confirmed that children were differentially likely to change their style of resource distribution across the two Fairness Types (Merit Intervention vs. Equality
Intervention), $b = 1.02$, $p = .047$. This predictor became slightly weaker (and non-significant) after adding Age in Months and Gender into the model, $b = 0.99$, $p = .058$, although neither of these demographic variables predicted children’s changes in resource distribution, $ps > .75$.

Overall, these results show that children’s fairness preferences are malleable, and can be robustly influenced through very brief testimony – particularly when equality is preached.

4.2.2. Justifications of Resource Distribution

As in Study 1, when children’s fairness preferences shifted, these shifts were frequently aligned with changes in the content of justifications (see Fig. 3 and Supplementary Materials). Additionally, children who shifted to favoring merit frequently invoked considerations of relative input, even though this form of justification was absent amongst children who divided by merit at pre-test, tentatively suggesting a change in reasoning that was produced by the testimony.

4.3. Discussion

In this study, direct testimony robustly led children to shift their preferences for distributing resources according to principles of merit or equality, indicating that social communication can powerfully impel children to revise their existing moral beliefs. Participants were very quick to trust and adopt the first experimenter’s claims despite their contradiction of participants’ preexisting fairness preferences. This tendency was robust, despite previous research indicating that children do not passively internalize everything they hear, particularly when they hold countervailing beliefs (see Koenig & Sabbagh, 2013). The abundance of children who changed their preferred pattern of resource distribution in Study 2 is especially remarkable given that this intervention was much briefer than the storybook in Study 1 and did not require children to repeat back what they had heard. It is likely that the testimony was so effective because it tied directly to the pre-test distribution task, and because the lesson was not obscured.
by the irrelevant details of a narrative. This study additionally uncovered an effect of Fairness Type, such that testimony advocating for equality was more powerful than testimony advocating for merit, indicating that some fairness beliefs are more easily swayed than others.

Taken together, Studies 1 and 2 suggest a major difference in the relative effectiveness of storybooks and testimony for changing children’s fairness preferences. In Study 1, children did not revise their moral beliefs based on the lessons conveyed in the metaphorical stories, whereas in Study 2, direct testimony was highly effective in changing children’s patterns of resource distribution. It is thus tempting to conclude that testimony, but not storybooks, constitute an effective format for conveying moral values to children. However, there were several differences between the storybooks in Study 1 and the testimony in Study 2 that prevent drawing strong conclusions from their comparison. One significant dissimilarity is that the storybooks required a more abstract mapping to the post-test measure, as their moral lessons occurred in context of a fictional society of beavers. As such, the ineffectiveness of the storybooks in Study 1 could have resulted from children’s difficulty in detecting the relevant similarities or in applying the lesson from a story about animals to a disparate real-world situation, rather than limitations of a storybook format per se (Daehler & Chen, 1993; Walker & Lombrozo, 2017).

5. Study 3

To investigate whether the disparate findings across the previous studies were likely due to content or due to format, Study 3 carefully matched and directly compared storybooks and testimony. The storybooks were rewritten to share a number of critical surface-level similarities with the target distribution task. First, they portrayed relatable, gender-matched children. Second, the plot centered around the division of prizes to two students in an art class, which is highly
reminiscent of the target distribution task. Third, the discrepancy in performance between these two characters in the story (completing 25% of a task vs. completing 75% of a task) was exactly matched between the storybooks and the distribution task. Therefore, in comparison with the Study 1 storybooks, transfer was greatly facilitated in these storybooks due to the heightened structure-mapping between the intervention and the dependent measure (Gentner, 1983).

The testimony was rewritten to match the storybooks in every way possible. In order to achieve a direct matching of content and source to the storybook, the testimony took the form of a video-recorded monologue by an art teacher describing how she divides prizes in her art class, and her description very closely followed the plot of the storybooks. Utilizing videotaped testimony reduced potential variability in the presentation of the testimony, and it also reduced the possibility of unconscious, contingent responding that is more likely in live contexts.

Finally, to ensure that we were truly measuring enduring moral learning, rather than transient situational effects, participants were given a second post-test after a 2- to 4-week delay.

5.1. Method

5.1.1. Participants

Participants were 77 six-, seven, and eight-year-old children (35 female; $M_{age} = 7.40$; $SD_{age} = 0.86$; 24–28 children per age) who were recruited from a small city in the northeastern United States via a participant database and tested in a college laboratory. Participants were primarily White and middle- to high-SES. Participants were randomly assigned to hear either a Storybook or videotaped Testimony. As in the previous studies, participants were assigned to the Equality or Merit conditions based on their pre-test distributions. An additional 22 participants were tested but were excluded from analyses because they gave more stickers to the lower-performing child ($n = 1$) or because they failed a comprehension check (Storybook conditions: $n$
= 9; Testimony conditions: \( n = 12 \)). Once again, a greater number of children needed to be tested in the Merit conditions before 16 usable participants were obtained in each condition.

To address the extent to which the effects of the interventions would persist over time, participants were asked to return approximately 2 to 4 weeks after their initial study session. A total of 68 non-excluded participants (88.31% of the sample) were able to return at this second timepoint (mean delay = 21 days; range = 12–45 days).

5.1.2. Materials and Procedure

The general procedure from the previous studies was retained in Study 3, with the exception of the details described below. Some additional, exploratory questions were also asked, and these are described in the Supplementary Materials.

5.1.2.1. Intervention. All storybooks and testimony were matched in length (508 words) and in content, such that the primary difference was whether their content was conveyed through illustrations and the perspective of a narrator (Storybooks) or through the videotaped first-person account of an art teacher (Testimony). Each format involved an art teacher giving prizes to children in her art class, either based on the students’ performance on a coloring task or based on a desire for egalitarianism. In the Merit versions, the teacher came to appreciate the value of dividing prizes based on merit after seeing one child work very hard while another goofed off. In the Equality versions, the teacher came to appreciate the value of dividing prizes evenly after one child didn’t finish coloring due to having old, broken colored pencils, while another child finished coloring quickly due to owning brand-new colored pencils. In all cases, the art teacher began dividing prizes in accordance with the participant’s pre-test performance and then realized that it was better to divide prizes in a different way (see Fig. 4 for excerpts). The storybooks
were read directly to the children, with ostensive cues such as pointing to retain their attention throughout, as in Study 1. The video testimony was played on an iPad.

---

**Equality:** “Mrs. Jones asks her students to listen and makes an announcement. She tells them: ‘A good way to divide up the prizes is for everyone to get the same amount of prizes. No matter how many rainbows everyone colors, whether it’s six or two or anything else, everyone would get four prizes each. That’s a good way of dividing things up because it’s not fair for some people to get more prizes just because it’s easier for them to do a better job. Everyone will be much happier if they get the same number of prizes as everyone else.’”

**Merit:** “Mrs. Jones asks her students to listen and makes an announcement. She tells them: ‘A good way to divide up the prizes is for hard-working people to get more prizes than less hard-working people. The more coloring everyone finishes, the more prizes they’d get; for example, if they color six rainbows, they’d get six prizes! That’s a good way of dividing things up because it’s not fair for everyone to get the same number of prizes even though some people do more to deserve them. Everyone will be much happier if they get the prizes they worked for.’”

**Equality:** “When I thought about that, I realized something… A good way to divide up the prizes is for everyone to get the same amount of prizes. No matter how many rainbows everyone colors, whether it’s six or two or anything else, everyone would get four prizes each. That’s a good way of dividing things up because it’s not fair for some people to get more prizes just because it’s easier for them to do a better job. Everyone will be much happier if they get the same number of prizes as everyone else.”

**Merit:** “When I thought about that, I realized something… A good way to divide up the prizes is for hard-working people to get more prizes than less hard-working people. The more coloring everyone finishes, the more prizes they’d get; for example, if they color six rainbows, they’d get six prizes! That’s a good way of dividing things up because it’s not fair for everyone to get the same number of prizes even though some people do more to deserve them. Everyone will be much happier if they get the prizes they worked for.”

---

**Fig. 4.** Excerpts from each of the interventions that were used in Study 3, including a sample illustration from the storybook and a still frame from the videotaped testimony.
5.1.2.2. Comprehension Questions. To ensure that any differences that might emerge between Storybooks and Testimony could not be due to differences in comprehension across the two formats, we excluded any participants who did not accurately report the primary details of the intervention (see Participants section above). Specifically, participants were asked what the art teacher thought was fair at the beginning of the story or video, and what she thought was fair at the end of the story or video (with forced choice response options for each). As per a preregistered decision, participants were required to answer both questions correctly in order to be retained in the final sample.

5.1.2.3. Delayed Post-test. This task remained structurally matched to the pre-test and post-test. The superficial qualities of the task were changed, such that it involved matching tiles depicting various yoga poses. The target children were displayed as having matched either 3 of 12 tiles or 9 of 12 tiles.

5.2. Results

5.2.1. Primary Results

Participants were coded as having changed (1) or not changed (0) their type of fairness distributions between equality and merit from pre-test to post-test, and also from pre-test to delayed post-test. One-sample binomial tests comparing against baseline rates of change (21.88%) demonstrated that, collapsed across conditions, the interventions reliably led to changes in children’s distribution patterns from pre-test to post-test, $p < .001$, with 46.75% change overall, and from pre-test to delayed post-test, $p < .001$, with 58.82% change (see Fig.
This significant change was found for each of the interventions preaching equality, both at the initial post-test (Equality Storybook: 62.50% change; Equality Testimony: 62.50% change) and at the delayed post-test (Equality Storybook: 66.67% change; Equality Testimony: 66.67% change), $ps < .001$. For each of the interventions preaching merit, significant changes did not occur at the immediate post-test (Merit Storybook: 32.00% change; $p = .228$; Merit Testimony: 40.00% change; $p = .059$), although they did at the delayed post-test (Merit Storybook: 45.00% change; $p = .025$; Merit Testimony: 61.11% change; $p < .001$), suggesting that children might take longer to process or internalize lessons about dividing resources in accordance with merit.

A logistic regression, including only participants who returned for the delayed post-test, was conducted to predict whether children were differentially likely to change their style of resource distribution across the two Fairness Types (Merit Intervention vs. Equality Intervention), the two kinds of Intervention (Storybooks vs. Testimony), the Timepoint (Immediate vs. Delayed), and the interaction of these variables. Replicating Study 2, changes in resource distribution differed across the two kinds of Fairness, $b = 1.50, p = .042$. Despite this difference becoming descriptively less pronounced at the delayed post-test, there were no significant effects of Timepoint, $b = 0.90, p = .190$, or of the interaction between Fairness Type and Timepoint, $b = -0.61, p = .551$. Changes in patterns of resource distribution did not differ across the two kinds of Intervention (Storybook vs. Testimony), $b = 0.88, p = .212$, and there were no significant two-way or three-way interactions involving this variable, $ps > .39$.

---

2 The 68 participants who returned for the second session had a 45.59% rate of change at the initial post-test, which is similar to that found in the overall sample.
Adding Age in Months and Gender to the model led the effect of Fairness Type to become non-significant, $b = 1.19, p = .121$. The effect of Gender was also non-significant, $b = 0.22, p = .549$, but Age marginally predicted change, $b = 0.44; p = .071$. As there were no other significant effects in this model, we removed all predictors aside from Age. This led Age to become a significant predictor, $b = 0.57, p = .010$. Further exploration of the data demonstrated that 8-year-olds (62.50% change at immediate post-test; 78.95% change at delayed post-test) were much more susceptible to the interventions than either 6-year-olds (35.71% change at immediate post-test; 48.15% change at delayed post-test) or 7-year-olds (44.00% change at immediate post-test; 54.55% change at delayed post-test), suggesting that our broadened age range in this study allowed us to detect developmental changes that were not present with only 6- and 7-year-olds.

5.2.2. Justifications of Resource Distribution

Once again, participants’ justifications tended to align with their distribution patterns. Only participants who were influenced by the intervention tended to change the content of their justifications during the post-test (see Fig. 3 and Supplementary Materials for details).

5.3. Discussion

Study 3 indicated that closely-matched storybooks and testimony are similarly effective in their power to change children’s preexisting fairness preferences. Specifically, Study 3 found that approximately half of children revised their tendencies toward equality-based or merit-based distributions after an intervention, regardless of whether the intervention consisted of a storybook or testimony. The changes found in children’s fairness preferences were higher than those found in the Baseline Study, particularly when equality was preached, demonstrating that
both of the intervention formats were reliably effective. Furthermore, these changes persisted after a delay of several weeks, indicating that children’s preference changes were not fleeting.

These findings demonstrate that both modes of social communication can be effective in altering children’s fairness preferences. As in Study 2, this was especially true when convincing children that equality-based distributions are preferable to merit-based distributions, again suggesting that equality is a potent principle that is more easily made more attractive than merit, possibly because it is a default preference that is easily re-invoked with encouragement.

Although Studies 1 and 2 indicated that storybooks and testimony differ in their effectiveness when stories involve allegorical tales about animals – as they often do – and when testimony is conveyed directly rather than via a video, Study 3 demonstrated that both formats of social communication can be highly effective in changing children’s resource distribution tendencies.

6. General Discussion

The present research investigated whether adults’ social communication can change children’s minds about the best way to distribute resources to third parties. Findings indicated that children’s fairness preferences are susceptible to the influences of storybooks or testimony preaching the benefits of either equality-based fairness or merit-based fairness, but only in certain circumstances. Study 1 indicated that typical storybooks (fictional stories featuring animals, thus requiring far transfer) were not effective in changing fairness preferences. In contrast, Study 2 found that short in-person testimony was extremely effective in changing fairness preferences, with the majority of children revising their preexisting preferences after this direct intervention. However, Study 3 clarified that these differences were not due to the format of the social communication itself, but were rather attributable to features related to the distance
of transfer that are typically – but not essentially – aligned with these formats. Specifically, Study 3 found that storybooks and videotaped testimony that were closely matched to each other, as well as being closely matched to the target distribution task, were effective in changing fairness preferences in approximately half of participants.

Contrary to the Piagetian beliefs that inspired much of the early work on children’s resource distribution tendencies (e.g., Damon, 1977; Hook & Cook, 1979), children are not limited to thinking about fairness in a specific way as a result of stage-based constraints. These findings confirm that shifts from favoring equality to favoring merit are not fully governed by cognitive maturation or autonomous construction from past experiences. Rather, these preference changes can be enculturated. Both forms of reasoning can coexist in the minds of young children and children can switch between distribution styles even after a brief intervention.

These results also indicate that it may be easier to change fairness preferences by encouraging children to revert to an earlier-emerging form of thinking than by encouraging them to adopt a typically later-developing disposition. Across both Studies 2 and 3, equality-based interventions were slightly more effective than merit-based interventions. This imbalance may be explained by the robust finding that equality is a potent “stable attractor” in the moral domain (Baumard et al., 2012; Chernyak & Sobel, 2016; McAuliffe, Blake, Steinbeis, & Warneken, 2017; Shaw & Olson, 2012), thus exerting a strong pull on fairness preferences that may hold greater sway than appeals to merit. Notably, however, the gap between equality interventions and merit interventions became less pronounced after a delay in Study 3. Although this effect was not significant, it tentatively suggests that children may merely be slower to adopt merit-based preferences. Similar increases in acquisition after a delay have been previously been found in cases of moral learning (Rushton, 1975) and in other domains (e.g., Vlach & Sandhofer, 2012).
The comparison of Studies 1 and 2 suggest that there are extreme differences in the effectiveness of the kinds of storybooks and testimony that resemble those used in everyday life. Because there were several key differences between the storybooks in Study 1 and the testimony in Study 2, we eliminated these discrepancies in Study 3. Notably, the Study 3 storybooks focused on humans rather than on animals. In Study 1, it is possible that children resisted learning or transferring lessons from stories that are about animals (see Strouse et al., 2018). However, although extreme anthropomorphism has been shown to be detrimental to young children’s learning (Ganea et al., 2014), including in the domain of fairness (Larsen et al., 2017), this does not seem to apply to the more moderate forms of anthropomorphism that were used in the Study 1 storybooks (Geerdts, 2016; Geerdts, Van de Walle, & LoBue, 2016). In addition, the exploratory tasks in Study 1 indicated that the transfer from animals to humans per se was not the primary factor preventing the storybooks from being effective. Rather, transfer was likely hindered by the difficulty in structure-mapping more generally (Gentner, 1983), perhaps in connection with the Study 1 stories depicting worlds that were irrelevant to participants’ lives (Richert & Smith, 2011; Walker et al., 2015; but see Hopkins & Weisberg, 2017). In Study 3, by creating storybooks with a plot that was more easily mapped onto the resource distribution task, children became much more likely to transfer the lessons that were preached by the stories.

The lengthy, videotaped testimony in Study 3 was less effective than the brief, direct testimony used in Study 2. One explanation for this decrease in effectiveness is that testimony conveyed in person may more easily yield change in fairness preferences. At least for young children, information that is conveyed through video is not as effectively transferred as information that is conveyed directly (see Barr, 2019). Forms of testimony that are specifically tailored to a particular case at hand, as in Study 2, may also be especially effective.
Overall, despite a stark difference between storybooks and testimony when comparing the findings of Studies 1 and 2, the differences between these formats of social communication dissolved when they were carefully matched on multiple dimensions. Importantly, then, these findings confirm that children are capable of learning moral lessons from storybooks. Previous research has found that, with the right kinds of scaffolding (e.g., being provided with commentary or being prompted to explain events in a story), children can extract abstract moral lessons rather than focusing on superficial features (Mares & Acosta, 2010; Walker & Lombrozo, 2017). Here, we have found that this is also true when stories are specifically written to elicit preference changes in a relevant concrete situation. Thus, in circumstances where direct transfer is strongly facilitated, stories can serve as effective cultural vehicles for conveying moral information to children and producing moral belief revision. However, stories are no more effective than the consistently reliable method of directly telling children a moral lesson.

7. Conclusion

Once children have adopted a particular moral belief, what can motivate them to change their minds? The present research has demonstrated that social communication is a powerful mechanism for inciting belief revision, as long as the messages that adults convey are easily transferable to particular cases of resource distribution. This research also examined whether storybooks constitute a privileged format for changing children’s minds. Across cultures, storybooks are ubiquitous devices for influencing children’s beliefs and preferences about moral issues, which has led to much popular speculation about the subversive or positive effects that stories have on children (e.g., Kohl, 1996). Despite this folk belief that stories harbor a special potency for molding children’s values, the present research suggests that storybooks do not exert
a uniquely powerful influence on children’s fairness preferences. Storybooks must be closely matched to the moral problem at hand in order to be effective in leading children to alter their fairness preferences – and even then, stories are no more effective than non-narrative testimony. While stories may be inimitably compelling as sources of immersive entertainment, it is apparently the moral lesson, rather than the story itself, which serves as the critical factor that can influence children’s preexisting tendencies to prefer a particular moral stance.
Acknowledgements

We thank Maya Raad, Julia Rose, Julie Finke, Elizabeth Abraham, Lulu Gomez, Mira Lerner, Nicole Kolmstetter, and Kelsey Krook for their assistance with data collection; Heather Greenebaum and Caroline Lawrence for coding the open-ended responses; Rachel Ansell for illustrating the Study 1 storybooks; and Izzy Buck for providing the videotaped Study 3 testimony and granting us permission to use her image in the paper. We are especially grateful to the Norwell Public Schools, the Boys & Girls Club of Boston, the New School of Lancaster, and local families for participating in the research.
References


Kanngiesser, P., & Warneken, F. (2012). Young children consider merit when sharing resources with others. PLOS One, 7(8), e43979. doi:10.1371/journal.pone.0043979


Smith, C. E., Blake, P. R., & Harris, P. L. (2013). I should but I won’t: Why young children endorse norms of fair sharing but do not follow them. *PLOS One, 8*(3), e59510. doi:10.1371/journal.pone.0059510.s001


Click here to download Supplementary Item: Supplementary Materials.pdf