CHILD DEVELOPMENT

Child Development, xxxx 2018, Volume 00, Number 0, Pages 1–13

Asking Children to "Be Helpers" Can Backfire After Setbacks

Emily Foster-Hanson, Andrei Cimpian, Rachel A. Leshin, and Marjorie Rhodes
New York University

Describing behaviors as reflecting categories (e.g., asking children to "be helpers") has been found to increase pro-social behavior. The present studies (N = 139, ages 4–5) tested whether such effects backfire if children experience setbacks while performing category-relevant actions. In Study 1, children were asked either to "be helpers" or "to help," and then pretended to complete a series of successful scenarios (e.g., pouring milk) and unsuccessful scenarios (e.g., spilling milk while trying to pour). After the unsuccessful trials, children asked to "be helpers" had more negative attitudes. In Study 2, asking children to "be helpers" impeded children's helping behavior after they experienced difficulties while trying to help. Implications for how category labels shape beliefs and behavior are discussed.

An important social-cognitive challenge young children face is determining which of their own behaviors reflect fundamental and stable aspects of their identities and which are more temporary and situational. For example, if children help out in class by putting away toys, they might see doing so as a reflection of a stable aspect of their own identity, or simply as a behavior that they are doing in a particular situation. Linguistic cues are a key feature of experience that shapes this process—children are more likely to view characteristics as fundamental and stable components of identity if they are labeled with nouns (e.g., "Rose is a carrot eater") than otherwise (e.g., "Rose eats carrots whenever she can;" Gelman & Heyman, 1999; see also Markman, 1989; Walton & Banaji, 2004; Waxman & Kosowski, 1990). Bryan, Master, and Walton (2014) suggested that this type of lexicalization can be harnessed to promote pro-sociality in early childhood—they found that children (ages 4–5) engaged in more helping behaviors when asked to "be helpers" instead of "to help," perhaps because the use of the category label (*helpers*) motivated children to engage in behaviors that would allow them to enact and maintain membership in this stable (and positive) identity category (see also Bryan, Adams, & Monin, 2013; Bryan, Walton, Rogers, & Dweck, 2011 for similar evidence among adults). The suggestion from this work—to use categorical language to motivate pro-social action in young children—received broad attention (e.g., Kris, 2015; Singh, 2014).

Yet, category labels can have negative consequences as well. For example, Cimpian, Arce, Markman, and Dweck (2007) provided children with praise in the form of a category label (e.g., "you are a good drawer") versus a behavioral description (e.g., "you did a good job drawing"). When children were told they were doing well, there were no differences based on the form of the praise. But, after children experienced a setback (e.g., an instance in which they were told that they had made a mistake in their drawing), children who had initially heard the praise in the form of a category label (e.g., "you are a good drawer") had lower self-evaluations and decreased persistence on the task (see also Kamins & Dweck, 1999). Thus, describing children's behavior with category labels may have detrimental—rather than beneficial—consequences once children experience setbacks (Dweck, 2000; Kamins & Dweck, 1999), perhaps because setbacks may lead children to question whether they are really members of the positively valued group at all. Setbacks and difficulties are common features of children's experience throughout development and into adulthood, so it is critical

We thank Heba Hajir, Talena Smith, Lauren Lucitt, and Kaajal Sagar for assistance with data collection, and the Children's Museum of Manhattan for participating in this research. Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute Of Child Health & Human Development of the National Institutes of Health under Award Number R01HD087672 (PI: Rhodes) and Award Number F31HD093431 (to Foster-Hanson). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Correspondence concerning this article should be addressed to Marjorie Rhodes, New York University, 6 Washington Place, rm. 301, New York, NY 10003. Electronic mail may be sent to marjorie.rhodes@nyu.edu.

© 2018 Society for Research in Child Development
All rights reserved. 0009-3920/2018/xxxx-xxxx
DOI: 10.1111/cdev.13147

for a comprehensive theory of how category labels affect children's helping to examine these effects in the context of difficulty.

The present studies test the possibility that asking children to "be helpers" instead of "to help" has negative consequences for children's beliefs and behaviors once they experience setbacks (e.g., if they break a toy while trying to clean up). Although this prediction is indirectly supported by previous work on praise, it is important to test it directly for several reasons. First, the negative consequences of categorical language have been shown for evaluative language such as praise (e.g., telling a child that they are a "good drawer" instead of telling them they "did a good job drawing") but not for more neutral language (e.g., telling a child to "be a drawer" instead of "to draw"). Thus, it is an open question whether the use of such category labels in the absence of evaluative content can also have detrimental consequences once children experience setbacks. This question is particularly important given the ubiquity of this type of input in young children's daily lives (e.g., children are frequently called "readers," asked to "be scientists," and so on).

Second, much of the work on the negative consequences of categorical language (in the context of praise) and on the influential role of setbacks (e.g., Cimpian et al., 2007; Kamins & Dweck, 1999; Schunk, 1983) has been in achievement-relevant domains (e.g., involving tasks children might do at school), where it is reasonable to assume that the possibility of failure, or the idea that not all children will succeed in all tasks, could be quite salient. For pro-social tasks like helping, Bryan et al. (2014) suggest that the possibility of failure is not as salient for children and therefore that categorical language might have more robustly positive consequences. From this perspective, while young children undoubtedly fail at many tasks on which they try to help (e.g., dropping piles of recently folded clothes, spilling milk while trying to pour it), perhaps children view having *intended* to help or *trying* to help as sufficient to claim membership in the positively valued category, regardless of the outcome. Thus, it is important to determine whether children are sensitive to the outcomes of their helping actions, and whether negative outcomes can have similarly detrimental consequences for subsequent beliefs and behaviors in this domain as they do for achievement.

Finally, whereas Bryan et al. (2014) examined actual helping behavior, the related work in the achievement domain has primarily examined children's attitudes and beliefs in the context of hypothetical failure (Cimpian et al., 2007; Kamins & Dweck, 1999; but see Mueller & Dweck, 1998). Thus, it is possible that categorical language might lead children to feel more negatively or report that they would be less likely to persist on imaginary future tasks, but that the opportunity to actually engage in a behavior to maintain (or reclaim) membership in the valued category motivates children nonetheless. In this case, categorical language might differentially affect various aspects of children's beliefs, attitudes, and behavior.

The present studies thus examine how children respond to setbacks after they have been asked to "be helpers" or "to help." Study 1 employed a method similar to the relevant work in the achievement literature by having children act out pretend scenarios and then testing for subsequent effects on their beliefs and attitudes. Study 2 extended this examination to actual helping behavior, demonstrating the real-life relevance of research on categorical language.

Study 1

Method

Participants

The sample included sixty-four 4- and 5-year-olds (32 male; M_{age} = 4.97 years, range = 4.03–5.94 years). We used a design that was closely modeled on Cimpian et al. (2007) and planned to test the same hypothesis as was tested there—that children who heard categorical language would have more negative task-relevant attitudes following setbacks. A power analysis using the effect size reported by Cimpian et al. (2007) suggested a minimum sample size of 28 participants per condition for 90% power. Of our sample whose parents' provided race and ethnicity information (19% did not), 50% were White, 8% were Asian or Asian-American, 16% were Black, and 8% were Multiracial; also, 27% of the sample (who reported this information) were Hispanic (of any race). Socioeconomic status information was not collected. Children were recruited from the Children's Museum of Manhattan and were tested in a private room at the museum. Four additional subjects were tested; two were excluded from analysis due to experimenter error, and two for disruptions that occurred during the testing session.

Procedure

Children were randomly assigned to one of two conditions, the "To Help" condition or the "Helper" condition. First, children heard the same introduction used in Bryan et al. (2014). The experimenter said:

> Some children choose to help [be helpers]. You could help [be a helper] when someone needs to pick things up, you could help [be a helper] when someone has a job to do, and you could help [be a helper] when someone needs help.

Following the short introduction, children played a "school game" in which they chose a puppet to represent themselves and acted out six scenarios with another puppet representing a teacher, handled by the experimenter (similar to Cimpian et al., 2007; Kamins & Dweck, 1999). In each scenario, the teacher puppet asked the child puppet to help with a task and the child was described as attempting to help. The six scenarios used were helping to (a) put away crayons after drawing, (b) put away blocks after playtime, (c) pour milk during snack time, (d) clean up after snack time, (e) put away trucks after playtime, and (f) put away paints after painting.

The first four scenarios described the child successfully helping with the task, after which they were thanked. For example, children heard:

> One day, you were getting ready to have some milk and cookies at school. Then Teacher Debbie says, "(Child's name), will you help me pour the milk?" and you say "OK, teacher." And so you start to pour the milk into the cups, and the milk is very heavy. You really want to help [be a helper], and so you fill the cups all the way up with milk. Then Teacher Debbie says, "Thank you for helping [being a helper]."

The last two scenarios described the child making a mistake, such as dropping crayons while trying to help clean up, after which the experimenter responded with a mildly negative, "Hmm, that didn't go so well," in order to highlight the failure, but did not offer direct criticism. For example, children heard:

> One day you were drawing with some crayons. Then Teacher Debbie says, "(Child's name), will you help me put these crayons away?" and you say "OK, teacher." And so you start to put the crayons away in the cup where they go, and there are a lot of them. You really want to help [be a helper], but you knock over the cup and all the crayons fall out on the table. Then Teacher Debbie says, "Hmm, that didn't go so well."

All children heard six scenarios, but the six scenarios were rotated between subjects so that all scenarios were presented in successful and unsuccessful versions (to different children). The scripts used for all scenarios are available at: https://osf.io/ch89j/.

Attitudes toward helping. Our key dependent measure assessed children's attitudes toward helping with five questions, which were all asked after the third (successful) trial, and then again after the last (unsuccessful) trial: "Do you like helping, or do you not like it?" "Did what happened in the [crayon] story make you feel happy or sad?" "Did what happened in the [crayon] story make you feel like you were good at helping or not good at helping?" "Did what happened in the [crayon] story make you feel like you were a good boy/girl or not a good boy/girl?" "Imagine you're at the park tomorrow and you see someone who looks like they might need some help. Would you help them, or let someone else help them?" All items were scored 1 = more positive attitudes toward helping, 0 = more negative attitudes toward helping. For each child, the name of the story referenced in the questions was replaced with whatever was the most recent story (e.g., "the milk story"); thus, the first time these questions were asked (after the third story), they referenced a successful scenario, whereas the second time they were asked (after the sixth story), they referenced an unsuccessful scenario. We report these data as the probabilities of giving positive responses across the five attitude questions ($\alpha = .74$).

After the first and second questions, children were also asked to rate their level of agreement with each item on a 6-point scale (e.g., "how much do you like it/not like it?"). We did not find these responses to be more informative than children's binary answers, however. For this reason, and to allow us to consider children's responses to all of the attitude questions together in a single composite, we did not consider these follow-up questions further.

Additional measures. For exploratory purposes, we also assessed two other components of children's beliefs. First, we tested whether the effects of categorical language might generalize to how children evaluate other people's behavior (e.g., whether the categorical language led children to view

others' antisocial actions as reflecting stable components of their identities). After each set of attitude measures, children were asked two questions involving new characters (the characters were matched to the child's gender), for example: "Imagine a new student is in your class, Sally. She steals your crayons, scribbles on your paper, and tears your stickers. Then she calls you names. Do you think she will always act this way?" (1 = no, 0 = yes). "Imagine a new student is in your class. You look over at her and see that she broke one of the toys during playtime. Does this mean that she is bad?" (1 = no, 0 = yes). These two items were totaled (α = .42; results were the same when considering items individually) and are reported such that higher probabilities reflect more flexible beliefs about others' actions. Second, during the second block of questions only, we tested whether the categorical language changed children's hypothetical persistence, by asking whether they would choose to try an activity they had succeeded at (scored "0") or one that they had failed at (scored "1") if they had the opportunity to help again, and two questions about what they might do next after each failure experience (e.g., "Think about the time that you tried to help pour the milk but you spilled it on the floor. What would you do now?"). Explanations were coded by two independent coders (99% agreement, disagreement resolved by first author) as "1" for mastery-oriented responses (e.g., "Clean it up") and "0" for helpless responses (e.g., "Go to bed"); these three questions were totaled and are reported as probabilities of mastery-oriented choices and explanations (α = .41; results were unchanged when considering items individually).

At the end of the study, children acted out successful versions of the scenarios in which they had previously been described as failing so that participants were not left with negative feelings about the task.

Results

Attitudes Toward Helping

To examine the effects of our language manipulation, we implemented a mixed-effects model using the lme4 package in R (Bates, Maechler, Bolker, & Walker, 2015). The fixed effects in this model were the main and interactive effects of condition (to help vs. helper) and time (before vs. after failure). Because the children's ages spanned 2 years, we also included child age as a continuous predictor in all analyses. In both Study 1 and Study 2, we explored the possibility of including Age × Condition interactions as well but did not find that these terms improved model fits and thus did not include them in the final models. The models included random intercepts for each participant. We used the "glmer" function to specify a binomial distribution (as the attitude measure was composed from five binary responses).

Overall, children generally expressed positive attitudes about their own helping abilities (M = .83). Consistent with Cimpian et al. (2007), before failure, there was no effect of condition on children's attitudes, $\beta = -0.41$, $SE = 0.37$, $z = -1.12$, $p = .26$, whereas after failure, children had more negative attitudes in the "Helper" condition than in the "To Help" condition, $\beta = 0.60$, $SE = 0.27$, $z = 2.21$, $p = .03$. The odds of expressing a positive (vs. negative) attitude in the "To Help" condition were 1.82 (CI = 1.08, 3.12) times as high as in the "Helper" condition. Also, overall, children's attitudes were more negative after failure than before, $\beta = -1.8$, $SE = 0.38$, $z = -4.73$, $p < .001$, and the interaction between condition and time was marginal, $\beta = 1.08$, $SE = 0.56$, $z = 1.93$, $p = .053$ (see Figure 1). Data and code for all analyses and figures are available at: https://osf.io/ch89j/.

Additional Measures

There were no effects of condition, age, or time on either the composites or any individual measures of children's predictions about a third party (attributions to stable traits, helper, M = .42, SE = 0.26, to help, M = .36, SE = 0.26), or on their

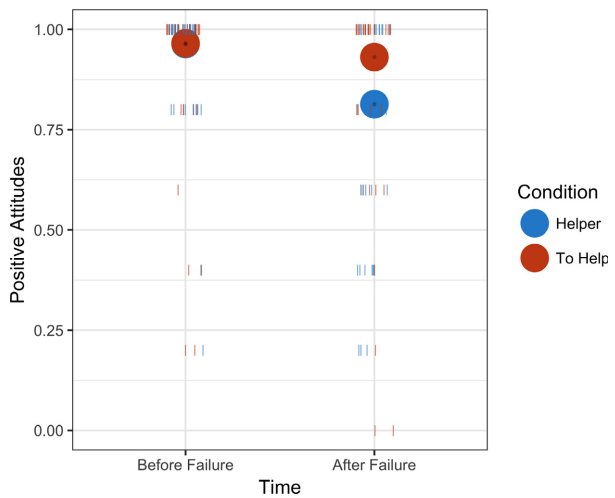

Figure 1. Probability of positive attitudes toward helping by condition, before and after failure; large dots represent means by condition, small lines represent individual children.

hypothetical persistence (probabilities of mastery-oriented responses, helper, $M = .77$, $SE = 0.24$, to help, $M = .78$, $SE = 0.25$).

Discussion

In Study 1, we found that the use of categorical language to describe helping undermines children's attitudes following setbacks. Children asked to "be helpers" had more negative attitudes toward helping than children who were asked "to help" after they experienced some difficulty with a helping task. In Bryan et al. (2014), wherein children did not experience a set-back, children were more motivated to help when they were asked to "be helpers" instead of "to help." Thus, this study suggests that category labels such as "helper" may backfire once children encounter a setback or experience difficulty.

We did not find an effect of condition on children's expectations of hypothetical persistence. The present studies examined children's attitudes following *pretend* scenarios, however, whereas Bryan et al. (2014) examined actual helping behavior. Thus, in Study 2 we examined whether asking children to "be helpers" influenced subsequent helping behavior after children experienced real setbacks brought about by their own actions.

In Study 1, we did not find that the effects of categorical language generalized to children's judgments of third parties. However, the behaviors queried in these measures asked about overall moral goodness rather than about helping specifically. In order for the language manipulation to have affected these judgments, children would have had to generalize not only from how they thought about themselves to how they thought about others, but also from how they thought about helping in particular to the social-moral domain more broadly. Therefore, in addition to examining children's helping behavior and attitudes, Study 2 also tested whether language affects children's beliefs about other people's helpfulness more specifically.

Study 2

The primary goal of Study 2 was to examine whether language influences children's actual helping behavior following experiences of setbacks. To do so, after implementing the language manipulation used in Study 1, we created situations in which children experienced two setbacks while trying to help, and then had three additional opportunities to help. These three opportunities to help were inspired by those used by Bryan et al. (2014)—asking children to put away blocks, open a container, and put away crayons—but our design and implementation differed from this previous research in several ways.

First, we examined children's behavior only after they experienced setbacks. We did not examine the effects of language on children's behavior in the absence of setbacks, as was done by Bryan et al. (2014), because of concerns about the overall length of the research session. Although this would have been useful, it was not central to our main research question, which was particularly about how language influences responses to setbacks. Also, to the extent that success scenarios serve to make the subsequent failure salient in hypothetical paradigms (such as Study 1), we did not think they would be needed here: The setbacks in Study 2 involved children's own behavior, so we expected them to be sufficiently salient to children on their own, even if they did not have earlier success experiences with which to compare them.

Second, although Bryan et al. (2014) did not examine their three helping tasks separately, in effect assuming that they are similarly indicative of whether children wanted to help, we reasoned that helping behaviors vary on a number of dimensions that could be relevant here. For example, such behaviors can differ in terms of how much effort they involve (e.g., picking up a few vs. a whole pile of blocks), how complex or challenging they are (e.g., picking up blocks vs. tying shoe laces), and whether the behavior benefits just the recipient or the helper as well (e.g., bringing over blocks for another child to play with vs. for the helper and recipient to play with together). By the preschool years, children consider many of these dimensions in determining whether to help. For example, they are sometimes less likely to help when doing so comes with more personal costs (Svetlova, Nichols, & Brownell, 2010), or when it requires more physical effort, even when the more effortful behaviors are within children's capabilities (Sommerville et al., 2018; see also Green, Kirby, & Nielsen, 2018). The more effortful and complex the helping behavior, and the more other-oriented (or selfless), the more it signals that a child has a strong motivation to help. As detailed next, the three tasks we used in Study 2 varied on these dimensions, thus providing a comprehensive test of the hypothesized backfiring effect of "helper" language on children's helping behavior. It is possible, for instance, that this effect would be present when the bar for helping is set

high (e.g., when the helping would require a lot of effort) but not when it is low. Thus, we decided in advance of data collection to include "task-type" as a within-subjects variable in our analysis to test for the possibility that the influence of categorical language on behavior following setbacks varies with the demands of the subsequent helping task.

Third, in addition to helping behavior, we also examined children's task-relevant beliefs and attitudes in order to connect the patterns of behavior in Study 2 with the self-report measures used in Study 1 and the broader literature on categorical language and setbacks (e.g., Cimpian et al., 2007; Kamins & Dweck, 1999).

Method

Participants

The sample included seventy-five 4- and 5-year-olds (36 male; M_{age} = 4.88 years; 63% White, 13% Asian or Asian-American, 6% Black, 18% Multiracial; 18% of the sample was Hispanic). Of the parents in our sample, 80% had a college degree or higher. Children were recruited from the Children's Museum of Manhattan and were tested in a private room at the museum. Fourteen additional participants began testing but were excluded from analyses: eight because they chose not to complete the study, three because of parent interference, and three because of mishaps with the study props (e.g., the false bottom did not fall out of the box on the first failure manipulation).

Procedure

Full scripts and pacing information for the experimental protocol are available at https://osf.io/ch89j/, and sample videos of the complete procedure in both conditions are available on Databrary (https://nyu.databrary.org/volume/481).

Language manipulation. As in Study 1, children were randomly assigned to one of two conditions, the "To Help" condition (n = 39, 19 male) or the "Helper" condition (n = 36, 17 male). Following a short warm-up activity (playing with play dough with Experimenter 1), children heard the same introduction used in Study 1 and by Bryan et al. (2014). In contrast with Study 1, however, children heard the language manipulation only during this initial portion of the study. To ensure that children understood the manipulation, the introduction was followed immediately by two manipulation check questions. Each question described a hypothetical scenario and asked the child to repeat back the condition manipulation (e.g., "Let's say the teacher was trying to pass out snacks to everyone. What could you do/be?" "Help/be a helper.") Participants who did not answer as expected immediately were prompted until they responded correctly (e.g., "You could help/be a helper, right? So, what could you do/be?"). Experimenter 1 conducted the condition manipulation portion of the study with the participant; Experimenter 2 was not present in the room during the condition manipulation and was thus blind to the condition to which the child had been assigned. Following the introductory phase just described, Experimenter 2 was brought into the room and conducted the remainder or the study with the child, whereas Experimenter 1 sat in a corner facing away from the child "to work on some papers."

Failure trials. After entering the room and introducing herself, Experimenter 2 engaged the child with a new activity (coloring with crayons). Once the child was engaged in the activity, the experimenter initiated the first of the two trials that were designed to elicit experiences of failure, by saying, "Oh, whoops! I forgot to put that box away on the table." If the child did not immediately begin to help to put the box away, Experimenter 2 provided the child with scripted prompts, one at a time, until the child began to help. The box that the child tried to help put away was full of ping pong balls and had a loose bottom. When the child attempted to pick the box up to put it away, the bottom fell out and the balls fell out onto the floor. The experimenter responded to this event by saying, "Uh oh. I guess I can do that later" in a neutral tone (an audio clip of the experimenter's response is available at https://osf.io/ch89j/), and then reengaged the child with coloring.

After a brief time, the experimenter then initiated the second trial that was intended to elicit failure by saying, "Oh whoops! I forgot to put that toy truck away in the toy bin." Again, several scripted prompts were used until the child began to help. On this trial, the toy truck that children attempted to put away had been previously disassembled into pieces and then carefully placed so as to look whole. When the child attempted to pick up the toy truck, it "broke" into several pieces. The experimenter responded with, "Uh oh. I guess I can do that later." All children helped on these two trials and thus experienced these setbacks. We used multiple prompts, including directly asking for the child's help, if necessary, to encourage helping, as

these experiences were an important part of the design.

Dependent measures. Experimenter 2 then reengaged the child with the coloring activity. Subsequently, children were exposed to three additional opportunities to help. For these trials, the experimenter's words and actions were intended to draw children's attention to opportunities to help (e.g., by saying "Oh whoops! I forgot to . . ."), but unlike for the two initial failure trials, the experimenter never explicitly asked for the child's help. Thus, we expected variation in children's helping behavior on these trials.

These three opportunities varied on a number of dimensions. Crucially, two of them required the child to disengage from what they were doing, walk to a different area of a room, and interact with props that were unfamiliar and irrelevant to them, all to help a research assistant with whom they were also unfamiliar. We thought that these situations would create stringent tests of children's motivation to help. In contrast, the third task did not require them to leave their seat, and involved picking up materials with which they themselves had just been playing, would continue to use, and in fact needed to continue their activity. Because of these differences, we made an a priori decision to analyze the tasks separately rather than simply summing across them. Consistent with this decision, a reliability analysis suggested that the three tasks had low internal consistency ($\alpha = .54$). Thus, we included task-type as a within-subjects factor in our model and tested how the language condition influenced children's behavior across the three different opportunities to help.

Task 1 consisted of a pile of blocks that needed to be put away into bags. The blocks were located on the other side of the room from where the child was sitting and were obscured from view by a table. As the experimenter put the blocks away, she drew the child's attention to the opportunity to help using a series of scripted verbal prompts (e.g., "Hmm, this is hard to do by myself.") Helping on this task would require the child to notice that the experimenter was engaged in a difficult task on the other side of the room, stop what they were doing at the time (coloring), leave their seat, and walk across the room to help clean up a mess that they had not been involved in creating (and involving a resource that was not for them to use). Helping on this task would benefit only the researcher, not the child. Also, because children had not interacted with these blocks before and just experienced two setbacks while trying to help with materials in the room (e.g., a toy that unexpectedly broke into pieces), we expected them to be somewhat uncertain of their ability to help with these new materials. For all of these reasons, we considered this to be an effortful helping task along the lines of the dimensions described earlier, and thus, one that is particularly susceptible to the consequences of the previous categorical language.

Task 2 involved lifting the lid off of a storage bin so that the experimenter could place a heavy stack of trays inside. Like Task 1, helping on this task required children to notice the experimenter's need for help, stop what they were doing, get up, and walk to another area of the room in order to help put away a resource that they had not used and did not expect to use. Thus, this helping behavior would be for the sole benefit of the experimenter. Moreover, given children's recent failure experiences and the fact that they had never handled the experimenter's trays before, it is also possible that children did not feel certain of their ability to succeed in this context. Additionally, the prompts used to draw attention to the opportunity to help on this task were subtler than those in Task 1 and predominantly nonverbal: While holding the stack of trays with both hands, the experimenter attempted to lift the lid with her foot repeatedly while saying, "Hmm;" no other verbal prompts were given on this task. Thus, children had to attend closely to the needs of the experimenter to complete this task. For all of these reasons, we also considered this to be an effortful helping scenario, which again would be particularly susceptible to the negative consequences of the categorical language.

Children were then given the opportunity to help with a third task, which provided an easier opportunity to engage in helpful behavior (and perhaps end the study on a positive note). For this task, the experimenter came over to where the child was sitting coloring with crayons and brought more "crayons for the coloring game." After placing them on the table, the experimenter reached over to point at some aspect of the child's drawing (e.g., "Wow, look at that blue sky!") and "accidentally" knocked the crayons onto the floor just below the child. The child then had the opportunity to bend over to help pick the crayons up. Note that (a) this task did not require the child to get up; (b) the crayons could be used immediately by the child for coloring; (c) the child had been at least passively involved in the creation of the mess; and (d) the child had interacted with the crayons before and knew there was nothing mysterious about them—all the task required was

bending over, picking the crayons up, and putting them back on the table. Thus, Task 3 was the least effortful and uncertain, and it also offered the most direct benefit to the child.

Attitudes and beliefs. After all of the helping trials, children were asked a series of questions to assess their attitudes and beliefs about helping. These included three self-evaluation questions used in Study 1, which asked how good the child was at helping, how much they liked helping, and how happy they were with what had happened. For each, children were asked a binary question (e.g., "Are you good at helping or not good at helping?"), and then a follow-up based on their initial response (e.g., "Are you a little good, pretty good, or very good at helping?"). As in Study 1, we did not find these follow-up responses to be more informative than children's binary answers; therefore, we did not consider them further. In order to test whether language affects children's beliefs about other people's helpfulness, we then asked children a series of four questions about another hypothetical child who had encountered a setback while trying to help. Specifically, we asked whether the character (a) was good at helping or not, (b) would try to help another day or not, (c) would succeed or fail at another helping task, and (d) whether the participating child would want to play with the character.

Coding

Research sessions were video recorded from three camera angles. Video data were then processed to separate the manipulation phase (with Experimenter 1) from the remainder of the testing session (with Experimenter 2) and coded using the Datavyu behavioral video coding software by a set of trained research assistants who were blind to the participant's condition throughout the coding process. A randomly selected 25% of participants for each construct (see following) were coded by a second set of independent coders for reliability. Inter-coder reliability was 95%. The complete coding manual is available on Databrary (available to authorized users with sign-up; https://nyu.databrary.org/volume/481); key codes will be summarized here.

Language manipulation. The manipulation segments of the videos were coded for fidelity to the experimental protocol, the amount of time spent administering the condition manipulation script, and children's initial responses to the manipulation check questions (coded "1" for correct and "0" for incorrect).

Helping behavior. Our key dependent measure was whether or not children helped on each of the three helping tasks (scored "1" for helping and "0" for not helping for each task). These data were available for all participants.

We coded additional measures for a subset of participants who had complete, usable video data from all three of the camera angles (which was not the case for all participants, due to camera problems, children stepping out of the camera range, and so on). For the participants with complete, usable video data (N = 49; n to help = 24, 12 male; n helper = 25, 11 male), we coded for several additional constructs: (a) emotional distress in response to the failure manipulation (rated by the research assistants on a scale: 0 = *not at all distressed*, 1 = *a little distressed*, 2 = *very distressed*), (b) number of items picked up on the first difficult helping task (the "Blocks" task) and the last, easy task (the "Crayons" task), (c) time spent helping on each task, and (d) delay before onset of helping. Due to the relatively small sample of participants for which full video data were available, analyses of these more detailed codes should be interpreted more cautiously, and are available at OSF (https://osf.io/ch89j/).

Attitudes and beliefs. Children's responses to the questions about their self- and other-oriented beliefs and attitudes were given verbally and recorded on paper by the experimenter. In addition, responses were recorded on camera so that they were confirmable from video regardless of whether all three camera angles were available; thus, we have these data for almost the complete sample (N = 73; two children refused to answer these questions). Children's responses to the three self-evaluation questions and beliefs about the other character were scored "0" for a negative response and "1" for a positive response.

Results

Manipulation Understanding

Overall, children's initial responses to the manipulation check questions indicated that they understood the manipulation (72% accurate, CI = 64, 79); responses did not vary by condition (to help, M = 73%; helper, M = 71%; $\beta = 0.73$, $SE = 3.72$, $z = 0.2$, $p = .84$) or age ($\beta = 1.22$, $SE = 3.14$, $z = 0.39$, $p = .7$). Analyses revealed similar patterns regardless of whether children who passed or failed these questions were included, and children who initially failed these questions were prompted until

they responded correctly. Therefore, all children were retained for analyses.

Helping Behavior

We first examined whether children's helping behavior varied across the three tasks as a function of whether they had been asked to "be a helper" or "to help" (the model also included age and a random intercept for each participant). Overall, helping increased with age, $\beta = 0.98$, $SE = 0.43$, $z = 2.3$, $p = .02$. There was a main effect of task, with the third easy task eliciting more help overall than the two effortful tasks ($\beta = 1.89$, $SE = 0.67$, $z = -2.84$, $p = .005$), consistent with our a priori expectation that this task would offer an easier opportunity for children to engage in helping behavior. There was also a significant interaction between condition and task, $\beta = -2.08$, $SE = 0.88$, $z = -2.36$, $p = .02$. As shown in Figure 2, children in the "To Help" condition tended to help more than children in the "Helper" condition across the two effortful tasks, $\chi^2(1) = 3.24$, $p = .07$, whereas this pattern reversed on the easy task, $\chi^2(1) = 2.27$, $p = .13$. This pattern is broadly consistent with the idea that children who had been told to "be helpers" but then made mistakes were overall less motivated to help than children who had been told to "help": When the helping tasks required a considerable amount of effort and also carried the possibility of another failure, the helper children were hesitant to help; it was only when helping was low-effort and success was likely that the helper children actually helped the experimenter.

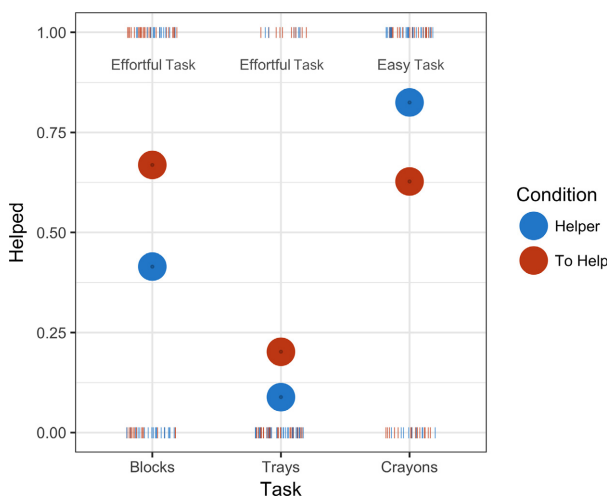

Figure 2. Probability of children's helping behavior by task and condition; large dots represent group means, small lines represent individual children. As reported in the main text, children in the helper condition helped less often on the first, effortful task than the last easy task, whereas children in the to help condition were as likely to help on these two different tasks. For comparisons of the second effortful task to the easy task, children in the helper condition also helped more often on the third easy task than on the second effortful task ($\beta = 4.20$, $SE = 1.09$, $z = 3.86$, $p < .001$). In this case, children in the To help condition showed a similar but less pronounced tendency ($\beta = 1.79$, $SE = 0.63$, $z = 2.81$, $p = .005$). As shown, overall rates of helping on the second task were quite low, presumably because it was a very challenging task (as described in the main text).

The condition differences in children's patterns of helping behavior across the tasks reinforce this point. Children in the to help condition were just as likely to help on the (first) effortful task as on the (last) easy task, $\beta = -0.18$, $SE = 0.55$, $z = -0.32$, $p = .75$, whereas children in the helper condition were *much* less likely to help on the (first) effortful task than the (later) easy task. In fact, the odds of helping on the easy task were 7.77 times as high (CI = 1.76, 34.39) as on the (first) difficult task for the helper children, $\beta = 2.05$, $SE = 0.76$, $z = 2.7$, $p = .01$.

Attitudes Toward Helping

We did not find a significant main effect of language condition on responses to the attitude measures included at the end of the study. By the time children got to the attitude measures, however, their experiences varied not only according to the condition manipulation, but also to the extent that some children had done more helping behaviors by this point than others. Thus, for exploratory purposes, we examined children's self-attitudes as a function of condition and whether they helped on at least one of the two effortful helping tasks (to help, $n = 27$, helper, $n = 18$) or whether they did not help on either of these tasks (to help, $n = 12$; helper, $n = 18$). We again used the "glmer" function with a binomial distribution, and tested as possible predictors the main and interactive effects of condition and whether children had ever helped on an effortful task. The model also included age as a covariate and a random intercept for each participant. This analysis revealed that the effect of condition on children's attitudes varied by whether they had helped on one of the effortful tasks or not, $\beta = 4.01$, $SE = 1.65$, $z = 2.43$, $p = .02$. Only among children who did not help, those in the "Helper" condition had lower self-evaluations than those in the "To Help" condition, $\beta = 2.53$, $SE = 1.25$, $z = 2.03$, $p = .04$ (see Figure 3). Given the exploratory nature and relatively small sample sizes per cell for this comparison, and the fact that this

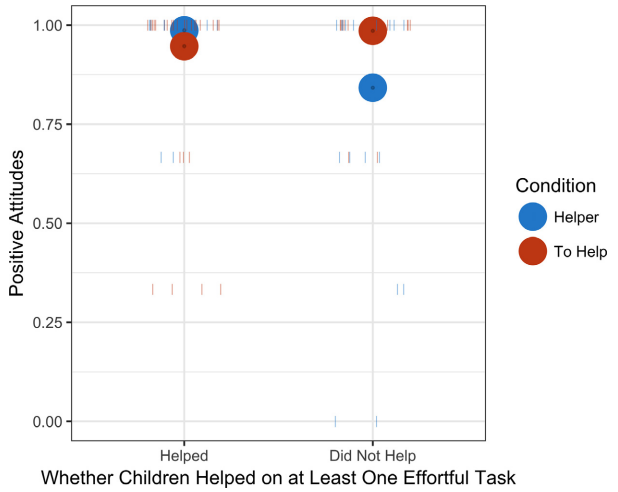

Figure 3. Probability of positive responses on self-attitude questions by condition, and whether or not child helped on at least one difficult helping task. Large dots represent group averages, lines reflect individual participants.

analysis included as a predictor children's previous choices about whether to help (which could also relate to other characteristics of the participants), this finding should be interpreted with caution.

There were no effects of condition or of children's own helping behavior on their judgments about the helpfulness of another child. Participants responded that that the other child was not good at helping ("good" = 1, "not good" = 0; M = .26, SD = .44), that he or should would try to help pour milk for snack if asked on another occasion ("would try" = 1, "would not try" = 0; M = .68, SD = .47), but that he or she would likely spill some of the milk ("would not spill" = 1, "would spill" = 0; M = .11, SD = .31), and said they would like to play with him or her if given the chance ("would play" = 1, "would not play" = 0; M = .79, SD = .41).

Discussion

In Study 2, children who had been asked to "be helpers" but then made mistakes helped less on difficult than on easy tasks, whereas children asked "to help" subsequently chose to continue helping regardless of task difficulty. Additionally, children asked to "be helpers" who subsequently chose not to help on either of the effortful tasks had lower self-evaluations of their helping abilities than children who had been asked "to help." These data indicate that categorical language can have detrimental consequences for children's behavior, even in nonacademic domains and even when the categorical input is not evaluative in content.

We do not deny that there may be circumstances in which being asked to be a helper is beneficial (e.g., children in the "helper" condition were somewhat more likely to help on the final, easy task of picking up crayons); however, what our results suggest is that those circumstances are more limited than previously thought, and that the benefits may be outweighed in some circumstances by the risks of this language backfiring. It is also important to note that we did not examine how language influences helping behavior in the absence of failure experiences (as was done by Bryan et al., 2014). Thus, the present data do not contradict those previously reported. Rather, we view the present studies as building on this earlier work to suggest a class of situations where categorical language might backfire.

Finally, note that we did not find evidence that categorical language affected judgments of another child in either Study 1 or Study 2. Thus, future research will be needed to determine when (if ever) children generalize from how they see themselves following categorical language to inform their beliefs about and expectations of other people.

General Discussion

These two studies reveal that, under certain circumstances, categorical language (asking children to "be a helper" instead of "to help") negatively influences children's self-evaluations and patterns of helping behavior after they experience setbacks while trying to be helpful. Despite the wide appeal of advice that encouraging children to "be helpers" can lead to increased helping behavior, the current data suggest that this type of categorical language can in fact backfire once children encounter the kinds of setbacks that are inevitable throughout early childhood. Further, these data show that children are sensitive to the negative outcomes of their own actions even when they have positive intentions, and even in domains that are not achievement-relevant.

We found that category labels can be problematic for children even when they are not evaluative in nature. Here, we did not talk about "good helpers" versus "being good at helping," as in the literature on praise, but simply about choosing to "be a helper" or "to help." This means that the possible negative consequences of category labels extend far beyond the language input identified in previous

work. Category labels are common in speech to young children (e.g., referring to children as "readers," "little artists," and so on); the present findings suggest the importance of further work to pin down exactly when and why such labels can have problematic consequences.

Limitations

Several limitations of the current design suggestion the need for future work to test the generalizability of the effects that we have documented here across different types of situations. For instance, the specific failure experiences used in this study were contrived, as required by our experimental design. The types of events we set up are similar to those children experience in their daily lives—a toy breaking while a child tries to put it away, knocking over more items while trying to clean them up, spilling milk while trying to pour it, and so on. Nevertheless, the failure experience that we created could have been more salient to children than a typical experience would be because it occurred in an interaction with a stranger and in an unusual context, and perhaps also because two such experiences occurred in quick succession, which might have been more dramatic than the helping failures children would typically experience in their daily lives. Also, in the present studies, the experimenter did not respond to the setbacks with positive encouragement. Although the experimenters did not provide direct criticism, in Study 1 they said, "That didn't go so well" (as in Cimpian et al., 2007) and in Study 2 they said, "Okay, I guess I can do that later." These responses were intended to draw children's attention to the reality that setbacks had occurred, whereas in children's daily lives, their difficulties might elicit more supportive responses from parents and teachers (although perhaps sometimes more frustrated responses as well). Thus, it is possible that these particular responses from the experimenters contributed to the negative consequences that we observed (perhaps by denying children the opportunity to correct their mistakes and encouraging more helpless responses). While we acknowledge these potential limitations, it is important to note that these features of our protocol (i.e., the features that may have been different from children's daily lives) were consistent across our two conditions, so their presence cannot in themselves explain why children responded more negatively in the "helper" than "to help" condition. Identifying the extent to which these factors are necessary for eliciting the negative consequences of categorical language, however, will be an important goal for future work aiming to pin down more precisely when categorical language might be helpful and when it is likely to backfire. We suspect that multiple factors might contribute to these processes (beyond whether children experience a setback or not).

Future Directions

It would be useful in future work to examine if there are ways of responding to failure that would eliminate the negative consequences of categorical language (and perhaps restore the benefits documented by Bryan et al., 2014). Might highly positive and supportive responses from parents facilitate children's perseverance in the face of difficulty when trying to help? It will also be particularly useful to examine naturalistic data (e.g., of the kind reported by Dahl et al., 2017) to examine children's helping behavior in their typical environments, and to test how parents and caregivers talk about helping with their children and how caregivers (and then children) respond to difficulties.

Future work should also examine how the effects of categorical language and setbacks vary across different types of categories. For example, broad, abstract categories (e.g., "good people") may be more robust to setbacks because category membership is not linked to any specific behavior, so it is less likely that children would interpret difficulty enacting those behaviors as a threat to category membership. If a specific behavior is linked to these types of categories, however—for example, through person praise (Kamins & Dweck, 1999)—then setbacks might be even more threatening because these broad categories are so meaningful for children's identities. In contrast, some categories are so narrow (e.g., "carrot-eaters," Gelman & Heyman, 1999) that membership in them is not particularly meaningful for children's identities. The category "helpers" lies somewhere in between—it is broad enough to be socially relevant, yet it is also directly linked to a specific set of behaviors. Indeed, the morphosyntactic form of the word "helper"—constructed from the verb "help" plus the suffix "-er"—indicates the strong dependence of category membership on helping behavior.

The present data do not directly shed light on the mechanisms by which categorical language influences children's beliefs and behavior, or on the representations that underlie these effects. One possibility is that children interpret their failure experiences as indicating that they are not members of

the category at all (or perhaps even that they are members of a "not helper" group). Another possibility is that setbacks leave children feeling concerned that they are somehow "bad" members of the helper category (e.g., that they are "poor helpers"), or that others might view them as such. Likewise, children could stop helping because they no longer see it as an identity-consistent action (e.g., if they now view themselves as "not-helpers") or because they are afraid of finding more evidence that they are (or leading others to view them as) poor helpers. Future research should examine the mechanisms underlying these effects in more detail.

How precisely language and failure experiences influence children's behavior might depend on the representations of "helping" that children bring to the task. Bryan et al. (2014) suggested that children might view holding the *intention to help* as sufficient to confirm membership in this valued category. For this reason, Bryan et al. expected that children's helping would be fairly robust to failure experiences in this domain (because a failed helpful action still reflects a positive intention to help). We did not find this to be the case, perhaps in part because young children focus more heavily on the *outcome* than the intention behind a behavior. Although even infants consider intentions in their social reasoning under some circumstances (Hamlin, Wynn, & Bloom, 2007; Warneken & Tomasello, 2006; Woodward, 2009), the role of intentions becomes more robust across development, and preschool-age children (like those tested here) often base social judgments on outcomes more than intentions (Cushman, Sheketoff, Wharton, & Carey, 2013; Nobes, Panagiotaki, & Pawson, 2009). If this is the case, then it is possible that children would be more immune to these failure experiences as they grow older and begin to value their own intentions to help more than the outcomes. Consistent feedback emphasizing effort (which is tied to one's intentions) in early childhood might help facilitate such a shift in children's attention from outcomes to intentions (Gunderson et al., 2013).

The present data do contain some hints of possible benefits of the categorical language, as suggested by Bryan et al. (2014; see also Bryan et al., 2011; Schunk, 1983). For example, in Study 2, many children in the "helper" condition chose to help on the final, easier task. We have proposed that the influence of the categorical language varied across the subsequent helping tasks because the tasks themselves varied in how much effort they required, whether they benefited the participating child themselves, and the extent to which children had reason to feel confident in their own capacity for success on the task. In this way, the influence of categorical language on subsequent helping behavior depends both on whether children experience setbacks or not (as indicated by a comparison of the present findings to those previously reported by Bryan et al.) and on the demands of the subsequent helping tasks (as revealed in the present analyses). In future work, it will be useful to clarify more precisely which features of particular helping situations underlie variation in children's responses to categorical language. Also, it is important to note that some children in the "helper" condition did choose to help on the more effortful tasks, and that those who did so subsequently reported just as positive self-evaluations as children in the "to help" condition. Thus, in future work, it will also be important to examine individual variation in how children respond to these circumstances, to understand why some children in the "helper" condition were more resilient than others.

Despite these outstanding questions, the present studies provide an important caveat to previous messages to parents and teachers about how to use language to encourage pro-sociality in early childhood. In particular, these studies suggest that categorical language can have negative consequences to the extent that children are likely to encounter setbacks. Difficulty is an inherent part of early childhood, so asking children to "be helpers" might result in more helping behavior in the short term, but backfire once children inevitably encounter setbacks, particularly with more challenging tasks. Phrasing requests in terms of actions, instead of identity categories, might encourage children to view instances of difficulty as opportunities to learn rather than negative judgments of character.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi:10.18637/jss.v067.i01.

Bryan, C. J., Adams, G. S., & Monin, B. (2013). When cheating would make you a cheater: Implicating the self prevents unethical behavior. *Journal of Experimental Psychology: General*, *142*, 1001–1005. https://doi.org/10.1037/a0030655

Bryan, C. J., Master, A., & Walton, G. M. (2014). "Helping" versus "being a helper": Invoking the self to increase helping in young children. *Child Development*, *85*, 1836–1842. https://doi.org/10.1111/cdev.12244

Bryan, C. J., Walton, G. M., Rogers, T., & Dweck, C. S. (2011). Motivating voter turnout by invoking the self. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 12653–12656. https://doi.org/10.1073/pnas.1103343108

Cimpian, A., Arce, H. M., Markman, E. M., & Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychological Science*, *18*, 314–316. https://doi.org/10.1111/j.1467-9280.2007.01896.x

Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition*, *127*, 6–21. https://doi.org/10.1016/j.cognition.2012.11.008

Dahl, A., Satlof-Bedrick, E. S., Hammond, S. I., Drummond, J. K., Waugh, W. E., & Brownell, C. A. (2017). Explicit scaffolding increases simple helping in younger infants. *Developmental Psychology*, *53*, 407. https://doi.org/10.1037/dev0000244

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. New York, NY: Psychology Press.

Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories. *Psychological Science*, *10*, 489–493. https://doi.org/10.1111/1467-9280.00194

Green, M., Kirby, J. N., & Nielsen, M. (2018). The cost of helping: An exploration of compassionate responding in children. *British Journal of Developmental Psychology*. https://doi.org/10.1111/bjdp.12252

Gunderson, E. A., Gripshover, S. J., Romero, C., Dweck, C. S., Goldin-Meadow, S., & Levine, S. C. (2013). Parent praise to 1-to 3-year-olds predicts children's motivational frameworks 5 years later. *Child Development*, *84*, 1526–1541. https://doi.org/10.1111/cdev.12064

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*, 557–559. https://doi.org/10.1038/nature06288

Kamins, M. L., & Dweck, C. S. (1999). Person versus process praise and criticism: Implications for contingent self-worth and coping. *Developmental Psychology*, *35*, 835–847. https://doi.org/10.1037/0012-1649.35.3.835

Kris, D. F. (2015, September 28). *Encouraging your child to become a "helper."* PBS Parents. Retrieved from http://www.pbs.org/parents/expert-tips-advice/2015/09/encouraging-child-become-helper/

Markman, E. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, *75*, 33–52. https://doi.org/10.1037/0022-3514.75.1.33

Nobes, G., Panagiotaki, G., & Pawson, C. (2009). The influence of negligence, intention, and outcome on children's moral judgments. *Journal of Experimental Child Psychology*, *104*, 382–397. https://doi.org/10.1016/j.cognition.2016.08.019

Schunk, D. H. (1983). Ability versus effort attributional feedback: Differential effects on self-efficacy and achievement. *Journal of Educational Psychology*, *75*, 848–856. https://doi.org/10.1037/0022-0663.75.6.848

Singh, M. (2014, April 30). *To get help from a little kid, ask the right way*. Shots: Health News from NPR. Retrieved from http://www.npr.org/sections/health-shots/2014/04/30/308045913/to-get-help-from-a-little-kid-ask-the-right-way

Sommerville, J. A., Enright, E. A., Horton, R. O., Lucca, K., Sitch, M. J., & Kirchner-Adelhart, S. (2018). Infants' prosocial behavior is governed by cost-benefit analyses. *Cognition*, *177*, 12–20. https://doi.org/10.1016/j.cognition.2018.03.021

Svetlova, M., Nichols, S. R., & Brownell, C. A. (2010). Toddlers' prosocial behavior: From instrumental to empathic to altruistic helping. *Child Development*, *81*, 1814–1827. https://doi.org/10.1111/j.1467-8624.2010.01512.x

Walton, G. M., & Banaji, M. R. (2004). Being what you say: The effect of essentialist linguistic labels on preferences. *Social Cognition*, *22*, 193–213. https://doi.org/10.1521/soco.22.2.193.35463

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*, 1301–1303. https://doi.org/10.1126/science.1121448

Waxman, S. R., & Kosowski, T. D. (1990). Nouns mark category relations: Toddlers' and preschoolers' word-learning biases. *Child Development*, *61*, 1461–1473. https://doi.org/10.2307/1130756

Woodward, A. L. (2009). Infants' grasp of others' intentions. *Current Directions in Psychological Science*, *18*, 53–57. https://doi.org/10.1111/j.1467-8721.2009.01605.x