Functional fixedness and cooties: Children solve insight problems faster when they learn functions from peers of a different gender

Herrissa Lamothe (hlamothe@mit.edu), Karla Perez (perezke@mit.edu), and Laura Schulz (lschulz@mit.edu)
Department of Brain and Cognitive Sciences, MIT, Cambridge MA 02139

Abstract

Knowing the intended use of an artifact impairs people's ability to think of alternative uses. Here we ask whether children consider not just "what" a tool is used for, but also "who" uses the tool in that way. We focused on gender roles since children are sensitive to these early in development and adapted a classic insight learning task (Defeyter & German, 2003). Success on the task requires inserting a stick into a tube to remove a ball. We compared children's (N = 112; ∼27 children/condition) latency to solve the problem at Baseline and three Demonstration conditions. In all the Demonstration conditions, the long stick was used as a magnet wand to brush away iron filings. In the Researcher Demonstration condition– a direct replication of Defeyter and German 2003– this function was demonstrated by a single individual – the experimenter; in the Same and Different Gender Peer Group conditions, the function was demonstrated by a group of children whose gender matched or differed from the participant's. Both Peer Group Demonstration conditions induced functional fixedness comparable to the Researcher Demonstration, and children were slower to solve the problem in all Demonstration conditions than Baseline. Critically however, children were faster to solve the problem in the Different Gender Peer Group condition than in the Same Gender Peer Group Condition, suggesting that children encode attributes of a function's typical user in their representations of artifacts, and that functional fixedness is affected by children's identification with a social role.

Keywords: gender; peer groups; functional fixedness; cognitive development

Introduction

Humans are sophisticated tool users (Heersmink, 2022), in part because we adopt a "design stance" (German & Johnson, 2002), inferring that artifacts have an intended or typical use. Knowing the intended use of an object inhibits our ability to think of alternative uses, a phenomenon known as functional fixedness (Defeyter & German, 2003; Dunker, 1945; McCaffrey, 2012; Kelem & Carey, 2006; Pflüger, Buttelmann, & Elsner, 2024; Ohlsson, 1984). Functional fixedness emerges cross-culturally, in both technologically rich and sparse cultures (Barrett, Laurence, & Margolis, 2008; German & Barrett, 2005), and although some work suggests that younger children may be immune to functional fixedness, it impairs insight problem solving in children age six and up (German & Defeyter, 2000). Here we ask whether children are sensitive not just to what an object is typically used for but who typically uses it that way. Are children sensitive to the gender of the person using the tool and does this affect their ability to imagine other ways of using the tool?

We focus on gender roles since they are pervasive in society and children are sensitive to gender-typed behavior early in development (Ellemers, 2018; Martin & Ruble, 2004). By the age of four, children prefer and attribute more positive traits to their own gender (Maccoby, 1998), selectively explore and have better recall for toys associated with their own gender (Bradbard, Martin, Endsley, & Halverson, 1986; Todd, Barry, & Thommessen, 2017), sanction peers who violate gender norms (Skočajić, Radosavljević, Okičić, Janković, & Žeželj, 2020; Sullivan, Moss-Racusin, Lopez, & Williams, 2018), and misremember counter-stereotype behavior as gender conforming (Martin & Halverson, 1983). Both adults and children also essentialize gender, minimizing variation between individuals, and treating gender-typed attributes as inherent, immutable, and due to membership in the gender category (Rhodes & Mandalaywala, 2017; Gelman & Taylor, 2014; Rhodes & Gelman, 2008, 2009; Taylor, 1996; Rhodes, 2014). Generic language is especially likely to elicit essentialist beliefs (Leslie, 2014; Cimpian & Markman, 2011; Rhodes, Leslie, Bianchi, & Chalik, 2018; Wodak, Leslie, & Rhodes, 2015) and it does so even when the generic language is intended to convey counter-stereotypical content (e.g., "girls are good at math"; (Benitez, Foster-Hanson, & Rhodes, 2024; Cimpian, Brandone, & Gelman, 2010), although contextual information can mitigate this effect (Vasilyeva, Gopnik, & Lombrozo, 2020).

Given children's sensitivity to information about gender, their readiness to associate particular toys and behaviors with gender roles, and their sensitivity to generic language, we hypothesized that children might be especially vulnerable to functional fixedness if they see a tool used in a particular way only by same-gender peers and especially immune from functional fixedness if the tool was used only by children of a different gender.

Experiment

Method

Participants Children were recruited from an urban children's museum and tested individually in private testing rooms off the museum floor. We tested five-, six-, seven- and eight-year-olds (N= 112; mean 7 years 0 months; range 4 years 11 months to 8 years 9 months). An additional 34 children were recruited but not included in the analysis; 11 were outside the age range; 14 were excluded due to experimenter error, 6 were excluded for touching the target items before the task was introduced, 2 were excluded for family interference, and 1 for missing birthday information with which to calculate age. Although we did not collect demographic information (beyond age and gender) for the participants, the population recruited for the study was representative of the

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

local population (48% Euro American, 22% African American, 20% Latino, 13% two or more races, and 10% Asian) and of museum visitors as a whole (40% receive free or discounted admission).

Children's parents were asked to complete a form reporting their child's preferred pronouns and gender identity (boy, girl, other, or prefer not to say) as part of the consent process. We used the parent report and did not independently ask the child to report their gender identity. Children were tested in the presence of their parents, and we did not want to put five to eight-year-olds in the position of possibly publicly contradicting their parents' identification. All children but one in our sample were identified by their parents as boys or girls; one child was identified as gender fluid and was assigned to a gender demonstration condition based on their gender identity that day.

Children were randomly assigned to one of four conditions: Baseline (N = 27), Researcher Demonstration (N = 27); Same Gender Peer Group Demonstration (N = 28) or Different Gender Peer Group Demonstration (N = 30), except that we tried to match the number of children in each age bin across conditions. Mean age did not vary significantly across condition (Mean Baseline: 6 years and 10 months; Mean Researcher Demonstration: 6 years and 9 months; Mean Same Gender Peer Group Demonstration: 7 years and 2 months; Mean Different Gender Peer Group Condition: 7 years and 2 months; $p = ns$ for all differences from Baseline). We performed a post-hoc power simulation to confirm our statistical model would recover the true value of a condition effect at 95% power with our existent sample size.

Materials *Demonstration items:* a (long) 32.1 cm magnet stick (a magnetic stir bar retriever) wrapped in either red or blue electrical tape and a 8"x8.5" custom-made black picture frame with iron filings covering a 4.5"x4.5" window depicting a picture under acrylic glazing. The custom-made black picture frame was covered with twelve glued on pom pom balls varying from 14mm to 20mm in diameter.
Distractor items: a pencil with both ends covered in masking tape (18cm total length) with a 17mm-diameter pom pom glued to one of the ends; a ruler wrapped in parchment printing paper (30.5cmx3cm) with a 20mm-diameter pom pom glued to one of the ends; and a (short) 15.4 cm magnet stick (magnetic stir bar retriever) wrapped in either red or blue electrical tape.
Task items: a clear round acrylic tube (length: 14", inside diameter: 17mm (0.67"), outside diameter: 20mm (0.8")). A glass marble (14mm (0.55")) wrapped in foam clay and then covered in silver metalized polyester mylar film tape (until it reached a diameter of 17mm and 16mm in length) to create a metal looking (but non-magnetic) bullet-shaped ball that would fit very snugly in the tube, getting well and truly stuck.
Gender Peer Group Videos and Pictures: prior to the study, 18 children (9 boys, 9 girls, ages 4- to 8-years old) in the same age group were recruited at the museum site and filmed while using the demonstration items. Four montages of all boy-presenting children or all girl-presenting children were created. These montages depict one 60 second video of a child using the tool for the entire duration of the video clip, and 8 still images of other children using the demonstration items. Each child in the montage was given a gendered pseudonym (e.g., 'Natalie' or 'John') which was displayed along with their image (see Figure 1). Two versions of each all boy or all girl montage was created, one depicting the boys/girls using the long blue magnet stick; and one depicting the boys/girls using the long red magnet stick. An iPad was used to show participants the videos and pictures of other children using the tool in the Gender Peer Group Demonstration conditions. All items were placed on a 28"x14" shoe tray positioned at the middle of the table. A second indistinguishable set including demonstration, distractor, and task items was constructed to allow for two researchers to conduct the experiments simultaneously in two separate experiment rooms.

Procedure In a direct replication of Defeyter and German (2003), children were told that they would be playing with items belonging to a puppet (here named "Kartik"), and that they would need to help Kartik as part of the game. Also replicating (Defeyter & German, 2003), the solution to the insight problem was always to use the long stick (counter-balanced red/blue) to poke the ball out of the tube; however, in the demonstration conditions, children always saw the long stick used first as a magnet wand to remove metal filings from a picture.

Researcher Demonstration Condition: In the researcher demonstration condition, children were shown how to use the long magnet stick to brush away the metal filings in the picture frame by a single adult demonstrator – the experimenter. They were told: "See, there is some metal dust hiding the picture. And here is Kartik's blue/red stick. It has a magnet on the end and Kartik uses it to move metal things. Look, here I am using it to brush away the metal dust hiding the picture". The demonstration by the experimenter lasted 30 seconds. Following the demonstration, the experimenter placed the two demonstration items on the shoe tray on the table, placing the magnet stick on top of the picture frame in order to prime its function, and brought out the distractor items and laid them out on the shoe tray, evenly spaced out. The experimenter then introduced a transparent tube with a silver (non-magnetic) marble stuck in the middle of the tube. Children were told, "now Kartik needs your help. Kartik went and got this metal ball stuck in this tube, and they need it for a magic show tonight. Can you help Kartik get it out? You can use any of Kartik's things, but one at a time, to get this metal ball out." Children were then handed the tube and began the insight task.
Gender Peer Group Demonstration Conditions: The two Gender Peer Group conditions were identical to each other, except that children saw either a group of children of their own gender, or a group of children of a different gender; and they were primed with both the object's function as well as

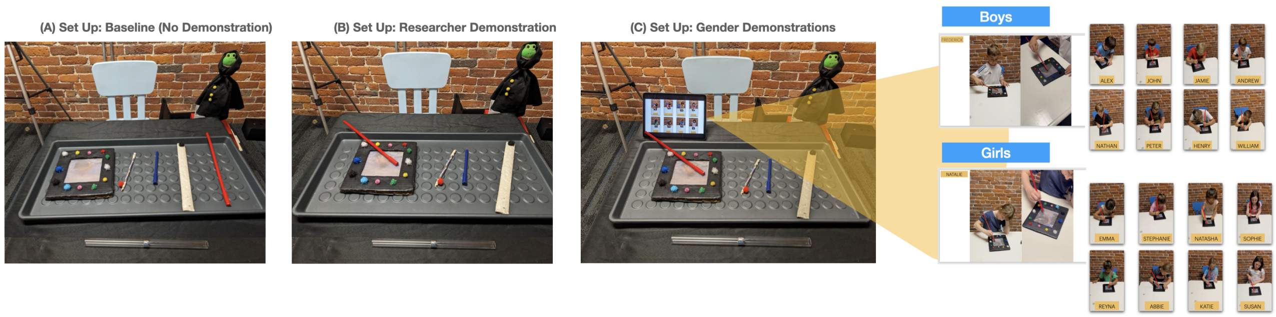

Figure 1: Experimental condition set up. (A) Baseline condition set up – all items are laid out on the shoe tray with no demonstration involved. (B) Researcher demonstration condition set up – function of long magnet stick is demonstrated by experimenter and magnet stick is left in a function prime position during test (on top of the picture frame). (C) Gender peer group demonstration condition set up – the magnet stick's brushing function is demonstrated by a video and 8 image stills of boys or girls depending on the child's gender and experimental condition (same or different gender experimental group). In addition to the function prime, the iPad depicting the collage of stills is left on the table behind the demonstration items as a gender prime at test. The color of long and short magnet sticks (red or blue) is counterbalanced, so is the side of the puppet relative to the picture frame.

with the gender of the group of peer demonstrators at test. Children were told that "some of the children in the museum have already gotten a chance to play with Kartik's things. Let's watch what they did". In both Gender Peer Group conditions, the demonstration was introduced with generic language ("Let's see what the girls/boys chose to do with this tool").

Depending on the participant's gender and condition assignment, children were shown either the all boy or all girl montage, introduced with generic language. They were told: "Let's watch the girls/boys. This is Natalie/John, and they went ahead and chose Kartik's blue/red stick which Kartik uses to move metal things because it has a magnet on the end. See, they are using it to brush away the metal dust hiding the picture". After watching the single boy/girl demonstrating the magnet stick for 60 seconds, they were then shown the picture stills one at a time during another 60-second window. The researcher pointed out the pseudonym of each child in the stills, the fact that they chose to use the long magnet stick, and repeated what the stick was for. At the end of the gender demonstration, the experimenter repeated generic language, "so, all the boys/girls used Kartik's blue/red stick to brush the metal dust away from the picture". The experimenter then brought out the distractor items and presented the children with the insight task as in the Researcher Demonstration condition; but in addition to the function prime, the experimenter left the iPad on the table behind the demonstration items with a collage of the 8 all boy or all girl stills showing on the screen as a gender prime.

Baseline Condition: In the baseline condition, children were presented the demonstration and distractor items one at a time without being shown how any of these items were to be used. The experimenter laid out picture frame followed by the items on the shoe tray one at a time in a random order, saying only, "Here is one of Kartik's things, and another one...". Then the experimenter presented the children with the same insight task as in the other conditions.

Results: Main Effects

All results were coded from video recordings by researchers blind to condition. We coded for latency to solution after children were presented with the insight problem. We also coded whether children chose the long magnet stick on their first reach for a tool, and whether children emulated the demonstration in their problem solving strategies and tried to use the long magnet stick as a magnet (sweeping it on the outside of the tube near the silver ball).

Latency We began timing the insight learning task from the moment the experimenter let go of the tube unless the child grabbed the tube or another item first, in which case the task began when the child first touched the item. The end of the task was determined in one of four ways: the moment the child inserted the long magnet stick into the tube and touched the silver ball; the moment the child extracted the ball using another method (e.g., shaking the short stick vigorously in the tube until the ball fell out); the moment the child said they wanted to stop the task, or after 25 minutes of trying unsuccessfully. In practice, almost all of the children were timed out using the first measure; 5 children succeeded using the short stick, 2 children gave up and did not complete the task, and one child persisted unsuccessfully for 25 minutes.

Latency to solution ranged from a minimum of 3 seconds, to a maximum of 25 minutes (mean: 1 minute 50 seconds, median: 13 seconds). To address the right skewed latency to completion data, we log transformed latency and fit a generalized linear model (GLM) with a gamma distribution, good for analyzing latency outcome data, with an inverse link function. This model performed well against other alternatives models, see Table 1. Lastly, we identified influential outliers using Cook's distance and removed outliers exceeding the conventional threshold of 4/n, where n is the sample size, dropping 5 outliers.

Table 1: Model Comparison using AIC

Model	Latency	df	AIC	Best
Gamma:Inverse Link	Log Transform	5	380.83	*
Linear Model	Log Transform	5	418.73	
Gamma:Log Link	Original Scale	5	1171.96	
Linear Model	Original Scale	3	1541.46	

Children solved the task rapidly at Baseline (mean solution time: 6.33 seconds). Replicating previous findings of functional fixedness, children were slower in each of the demonstration conditions than at Baseline (all p-values compared to Baseline $< .001$).

As predicted, there was also a significant effect of who demonstrated the function on children's vulnerability to functional fixedness. Although children in the Different Gender Peer Group condition were slower to solve the task than children at Baseline (mean: 16.82 seconds vs. 6.33 seconds), they were much faster to solve the task than children in the Same Gender Peer Group condition (mean: 48.61 seconds; $p < .01$). This is not because children in the Same Gender Peer Group condition were more than usually vulnerable to functional fixedness. Although children in the Same Gender Peer Group condition were slower than children in the Researcher Demo condition (mean: 32.42 seconds) the difference did not reach significance ($p = ns$). Similarly, although children in the Different Gender Peer Group condition were faster to overcome functional fixedness than children in the Researcher Demo condition, this difference was only a trend ($p = .07$).

First Choice Tool We also analyzed children's first choice of tool in their problem solving strategies. We fit a logistic regression (GLM with a binomial distribution and logit link function) conventionally used for binary outcome data. At Baseline, the probability of choosing the long magnet stick (the correct tool to solve the problem) on first reach was 89% ($p < 0.001$). Children were less likely to choose the long magnet stick first in all the demonstration conditions compared to Baseline ($ps < 0.01$); none of the demonstration conditions differed significantly from each other (all comparisons $p = ns$).

Brushing Action We fit a logistic regression to analyze if children imitated the function demonstrated to them (a brushing action to move the iron filings) at test. Children in the Different Gender Peer Group condition were the only group of children who did not differ statistically from Baseline in their tendency to use the tool as a magnet, sweeping it on the outside of the tube (Baseline: 4% of children vs. Researcher Demo: 26%, $p < .05$; vs. Same Gender Peer Group: 29%, $p < .05$; and vs. Different Gender Peer Group: 13%, $p = ns$). This suggests that children were less influenced by the demonstration when it was provided by a group of different gendered peers than when it was provided either by the researcher or by a group of same gendered peers.

General Discussion

Taken together, these results suggest that children's artifact concept encodes not only what a tool is used for but who uses it that way. Any demonstration of a function of a tool impedes children's ability to imagine alternative uses, but the impediment (and tendency to imitate the demonstrated function) is much less when children see the function demonstrated by a group of peers of a different gender than a group of peers of the same gender. When children identify less with the typical user of the tool (the people the tool is for), they are better able to imagine other ways the tool can be used. Here, we used gender as the basis for children's social identification, but future work could look at whether the same effect of social distance on functional fixedness applies to other social concepts (age, race, language, etc.).

In the current study, we failed to replicate previous findings of an effect of age on children's sensitivity to functional fixedness: the five-year-olds were as vulnerable to functional fixedness in the demonstration conditions as the older children. Possibly the particular function we demonstrated (using the magnet wand) was more compelling to children than the functions demonstrated in previous work (e.g., a pencil used for drawing or a wand presented as a tool for making light), or possibly the particular population of children we worked with (visitors to an children's museum) had more robust artifact concepts. Future work might extend the study to younger children (three and four-year-olds) to see if immunity to functional fixedness does emerge, just at earlier ages.

Also, while we found that children were less vulnerable to functional fixedness given demonstrations by a group of peers of a different gender than by a group of peers of the same children, we did not find that the gender demonstrations made children either unusually vulnerable or unusually resilient to functional fixedness compared to when the demonstrations were carried out by an individual researcher. Nonetheless, there was a trend for children to be faster to solve the problem in the Different Condition than the Researcher demonstration, children in the Same Gender condition were the slowest to solve the task, and this effect was most pronounced in the oldest children. Future work might extend the study to older children (nine and ten-year-olds) to see whether they become more flexible about gender roles with age (and thus show no differential effects of gender demonstration on this task) or whether they increasingly associate artifact functions with gender roles with age, and thus become increasingly vulnerable to functional fixedness given demonstrations by same gender peers.

This study undertook a replication of the classic study De-

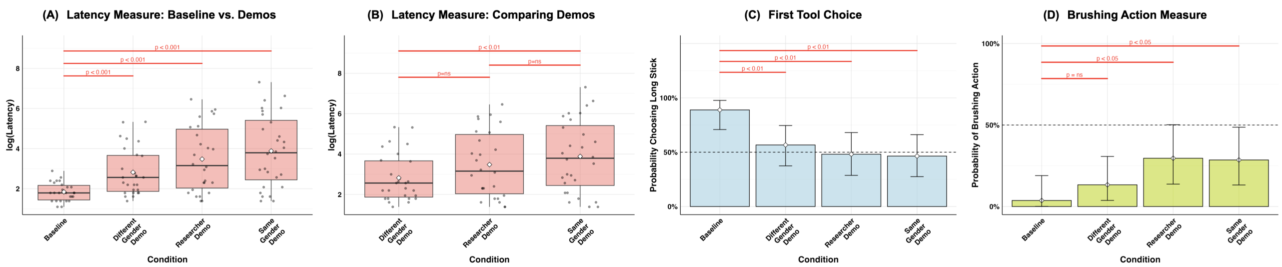

Figure 2: (A) Analysis of latency measure (logged) comparing all demonstration conditions to baseline; (B) analysis of logged latency measure comparing all three demonstration conditions; (C) analysis of first tool choice at test; (D) and analysis of imitation of brushing action (demonstrated function of magnet stick) at test.

feyter and German 2003, but focused on comparing two social learning conditions in which children were given demonstrations by a group of same or different gender peers. Future work might compare the effect of a single peer demonstrator to that of a group of peer demonstrators to see whether the gender of even a single peer demonstrator affects children's insight into alternative functions. Future work might also look at whether girls or boys as a whole are more affected by associating artifact functions with gender roles.

This study does not distinguish between universality and typicality of the tool users' genders. In this study, participants in the Gender Peer Group Demonstration conditions saw a group of all boys, or a group of all girls using the tool as a magnet wand. Future work might show children a majority (but not all) peers of the same gender using the tool as a magnet wand and establish whether children's insight learning is impacted only by seeing all same or different gendered peers, or whether it is also impacted by seeing mostly same or different gendered peers engaged in a behavior. Finally, this study looked only at girls and boys. Some work suggests that transgender children are as likely as cisgender children to essentialize gender roles (Gülgöz, Alonso, Olson, & Gelman, 2021; Gülgöz, DeMeules, Gelman, & Olson, 2019), suggesting that transgender children may also be as likely as cisgender children to incorporate gender roles into their artifact concepts, but future research might look at whether this is indeed the case, and whether children who are less likely to make binary distinctions between genders are less influenced by gender roles in inferring artifact functions.

The results so far, however, suggest that gender has a pervasive influence on the ease with which we see alternative possibilities, even for the use of simple tools. Children associated artifacts with gender roles given minimal evidence: just a few sentences of generic language, and two minutes of demonstrations. When children saw the demonstration by peers of the same gender, it took them more than twice as long to discover an alternative use of the tool than when the demonstration was by peers of a different gender. A little social distance appears to let us see things in a whole new way.

Acknowledgments

We thank our children participants at the museum and their families for making this study possible. Sienna Radifera and Nicole Coates for their assistance as well as all of the ECCL lab for their feedback.

References

Barrett, H. C., Laurence, S., & Margolis, E. (2008, January). Artifacts and Original Intent: A Cross-Cultural Perspective on the Design Stance. (Publisher: Brill) doi: 10.1163/156770908X289189

Benitez, J., Foster-Hanson, E., & Rhodes, M. (2024). *Generic references to gender predict essentialism even when they express counter-stereotypic ideas.* OSF.

Bradbard, M. R., Martin, C. L., Endsley, R. C., & Halverson, C. F. (1986). Influence of sex stereotypes on children's exploration and memory: A competence versus performance distinction. *Developmental Psychology*, *22*(4), 481–486. (Place: US Publisher: American Psychological Association) doi: 10.1037/0012-1649.22.4.481

Cimpian, A., Brandone, A. C., & Gelman, S. A. (2010). Generic statements require little evidence for acceptance but have powerful implications. *Cognitive science*, *34*(8), 1452–1482.

Cimpian, A., & Markman, E. M. (2011). The generic/nongeneric distinction influences how children interpret new information about social others. *Child development*, *82*(2), 471–492.

Defeyter, M. A., & German, T. P. (2003, September). Acquiring an understanding of design: evidence from children's insight problem solving. *Cognition*, *89*(2), 133–155. doi: 10.1016/S0010-0277(03)00098-2

Dunker, K. (1945). On problem-solving. *Psychological Monograph*, *58*(5).

Ellemers, N. (2018, January). Gender Stereotypes. *Annual Review of Psychology*, *69*(Volume 69, 2018), 275–298. (Publisher: Annual Reviews) doi: 10.1146/annurev-psych-122216-011719

Gelman, S. A., & Taylor, M. G. (2014). Gender essentialism in cognitive development. In *Toward a feminist developmental psychology* (pp. 169–190). Routledge.

German, T. P., & Barrett, H. C. (2005, January). Functional Fixedness in a Technologically Sparse Culture. *Psychological Science*, *16*(1), 1–5. (Publisher: SAGE Publications Inc) doi: 10.1111/j.0956-7976.2005.00771.x

German, T. P., & Defeyter, M. A. (2000, December). Immunity to functional fixedness in young children. *Psychonomic Bulletin & Review*, *7*(4), 707–712. doi: 10.3758/BF03213010

German, T. P., & Johnson, S. C. (2002, August). Function and the Origins of the Design Stance. *Journal of Cognition and Development*, *3*(3), 279–300. (Publisher: Routledge _eprint: https://doi.org/10.1207/S15327647JCD0303_2) doi: 10.1207/S15327647JCD0303$_2$

Gülgöz, S., Alonso, D. J., Olson, K. R., & Gelman, S. A. (2021). Transgender and cisgender children's essentialist beliefs about sex and gender identity. *Developmental science*, *24*(6), e13115.

Gülgöz, S., DeMeules, M., Gelman, S. A., & Olson, K. R. (2019). Gender essentialism in transgender and cisgender children. *PloS one*, *14*(11), e0224321.

Heersmink, R. (2022, October). Human uniqueness in using tools and artifacts: flexibility, variety, complexity. *Synthese*, *200*(6), 442. Retrieved 2025-05-09, from https://doi.org/10.1007/s11229-022-03892-8 doi: 10.1007/s11229-022-03892-8

Kelem, D., & Carey, S. (2006). The essence of artifacts: Developing the design stance. In *Creations of the mind: Theories of artifacts and their representation* (p. 212).

Leslie, S.-J. (2014). Carving up the social world with generics. *Oxford studies in experimental philosophy*, *1*.

Maccoby, E. E. (1998). *The Two Sexes: Growing Up Apart, Coming Together*. Harvard University Press. (Google-Books-ID: chtFzHfmBkMC)

Martin, C. L., & Halverson, C. F. (1983). The Effects of Sex-Typing Schemas on Young Children's Memory. *Child Development*, *54*(3), 563–574. (Publisher: [Wiley, Society for Research in Child Development]) doi: 10.2307/1130043

Martin, C. L., & Ruble, D. (2004). Children's Search for Gender Cues: Cognitive Perspectives on Gender Development. *Current directions in psychological science*, *13*(2), 67–70.

McCaffrey, T. (2012). Innovation Relies on the Obscure: A Key to Overcoming the Classic Problem of Functional Fixedness - Tony McCaffrey, 2012. *Psychological Science*, *23*(3). doi: https://doi-org.libproxy.mit.edu/10.1177/0956797611429580

Ohlsson, S. (1984). Restructuring revisited. *Scandinavian Journal of Psychology*, *25*(2), 117–129. (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9450.1984.tb01005.x) doi: 10.1111/j.1467-9450.1984.tb01005.x

Pflüger, M., Buttelmann, D., & Elsner, B. (2024, October). How children come to (not) detect and apply multiple functions for objects: Rethinking perseveration and functional fixedness. *Cognition*, *251*, 105902. doi: 10.1016/j.cognition.2024.105902

Rhodes, M. (2014). Inherence-based views of social categories. *Behavioral and Brain Sciences*, *37*(5), 501.

Rhodes, M., & Gelman, S. A. (2008). Categories Influence Predictions About Individual Consistency. *Child Development*, *79*(5), 1270–1287. (_eprint: https://srcd.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8624.2008.01188.x) doi: 10.1111/j.1467-8624.2008.01188.x

Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive psychology*, *59*(3), 244–274.

Rhodes, M., Leslie, S.-J., Bianchi, L., & Chalik, L. (2018). The role of generic language in the early development of social categorization. *Child Development*, *89*(1), 148–155.

Rhodes, M., & Mandalaywala, T. M. (2017, July). The development and developmental consequences of social essentialism. *WIREs Cognitive Science*, *8*(4), e1437. doi: 10.1002/wcs.1437

Skočajić, M. M., Radosavljević, J. G., Okičić, M. G., Janković, I. O., & Žeželj, I. L. (2020, February). Boys Just Don't! Gender Stereotyping and Sanctioning of Counter-Stereotypical Behavior in Preschoolers. *Sex Roles*, *82*(3), 163–172. doi: 10.1007/s11199-019-01051-x

Sullivan, J., Moss-Racusin, C., Lopez, M., & Williams, K. (2018, April). Backlash against gender stereotype-violating preschool children. *PLOS ONE*, *13*(4), e0195503. (Publisher: Public Library of Science) doi: 10.1371/journal.pone.0195503

Taylor, M. G. (1996). The Development of Children's Beliefs about Social and Biological Aspects of Gender Differences. , *67*(4), 1555–1571.

Todd, B. K., Barry, J. A., & Thommessen, S. A. O. (2017). Preferences for 'Gender-typed' Toys in Boys and Girls Aged 9 to 32 Months. *Infant and Child Development*, *26*(3), e1986. (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/icd.1986) doi: 10.1002/icd.1986

Vasilyeva, N., Gopnik, A., & Lombrozo, T. (2020). When generic language does not promote psychological essentialism. In *Proceedings of the 42nd annual conference of the cognitive science society.*

Wodak, D., Leslie, S.-J., & Rhodes, M. (2015). What a loaded generalization: Generics and social cognition. *Philosophy Compass*, *10*(9), 625–635.