Received: 30 October 2018 | Revised: 5 June 2019 | Accepted: 26 June 2019
DOI: 10.1111/desc.12885

PAPER

Developmental Science WILEY

Atypical shape bias and categorization in autism: Evidence from children and computational simulations

Ángel Eugenio Tovar | Angélica Rodríguez-Granados | Natalia Arias-Trejo

Facultad de Psicología, Universidad Nacional Autónoma de México, México City, México

Correspondence
Ángel Eugenio Tovar, Facultad de Psicología, Universidad Nacional Autónoma de México, Av. Universidad 3004, México City 04510, México.
Email: aetovar@unam.mx

Funding information
CONACYT, México, Grant/Award Number: CB 285152; DGAPA-PAPIIT-IN304417

Abstract
The shape bias, a preference for mapping new word labels onto the shape rather than the color or texture of referents, has been postulated as a word-learning mechanism. Previous research has shown deficits in the shape bias in children with autism even though they acquire sizeable lexicons. While previous explanations have suggested the atypical use of color for label extension in individuals with autism, we hypothesize an atypical mapping of novel labels to novel objects, regardless of the physical properties of the objects. In Experiment 1, we demonstrate this phenomenon in some individuals with autism, but the novelty of objects only partially explains their lack of shape bias. In a second experiment, we present a computational model that provides a developmental account of the shape bias in typically developing children and in those with autism. This model is based on theories of neurological dysfunctions in autism, and it integrates theoretical and empirical findings in the literature of categorization, word learning, and the shape bias. The model replicates the pattern of results of our first experiment and shows how individuals with autism are more likely to categorize experimental objects together on the basis of their novelty. It also provides insights into possible mechanisms by which children with autism learn new words, and why their word referents may be idiosyncratic. Our model highlights a developmental approach to autism that emphasizes deficient representations of categories underlying an impaired shape bias.

KEYWORDS
autism, categorization, computational model, shape bias, word learning

1 | INTRODUCTION

The shape bias is a preference for mapping new word labels onto the shape of referents, rather than on their color or texture (Hupp, 2015; Samuelson, 2002; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002). It has been studied mainly in English-speaking children. The word 'cup', for example, is efficiently extended to other objects whose shape resembles a child's cup, but it is not extended to objects that share the same color or texture. The shape bias is at the intersection of two cognitive processes: categorization and word learning. Recently, the analysis of the shape bias in children with autism spectrum disorders (ASD) has attracted considerable attention, as there has been no clear evidence of the shape bias in this population even when such children have acquired sizeable lexicons (Hartley & Allen, 2014; Potrzeba, Fein, & Naigles, 2015; Tek, Jaffery, Fein, & Naigles, 2008). This observation has raised important theoretical questions with far-reaching implications for therapeutic interventions in ASD: Are children with ASD using other strategies for word learning? Is the lack of a shape bias the result of atypical visual categorization and prototype formation in ASD? Are shape and other properties such as color of objects equally weighted for label extension in ASD? We argue that previous studies in this field have overlooked a common property shared by objects used in their experimental tasks: they are all novel objects. Our hypothesis is that children with ASD may be atypically classifying the experimental objects based primarily on their novelty, with color, texture and other object properties having a lesser effect on object categorization and consequently on label extension. We test this hypothesis in an empirical study with children and put forward a computational model that incorporates notions from neurodevelopmental theories of autism, categorization, and word learning to account for empirical data. Our objective is to provide a comprehensive approach to understand the typical and atypical development of the shape bias.

Research Highlights

- On average, children with autism did not show a shape bias for object names in an object-sorting task.
- We present a computational model of categorization and word learning in autism.
- Altered neural connectivity in a computational model of autism results in atypical categorization and an impaired shape bias.
- The model advances our understanding of the developmental trajectory of categorization and word learning.

ASDs are a group of neurodevelopmental disorders that frequently result in dysfunctional social interaction and communication, repetitive or disruptive behaviors, and narrow interests. Remarkably, the number and severity of symptoms are highly variable across individuals in the autism spectrum (Hu & Steinberg, 2009; Wozniak, Leezenbaum, Northrup, West, & Iverson, 2016), making it difficult to understand and characterize their cognitive and behavioral profiles. This heterogeneity is a major problem in developing a unifying theory of autism (Markram, Rinaldi, & Markram, 2007).

Studies of categorization abilities in persons with ASD have reported contrasting results. A number of studies have found the use of prototypes during generalization tasks to be impaired, either for natural stimuli like drawings of faces (Gastgeb, Rump, Best, Minshew, & Strauss, 2009), more artificial stimuli created from dot patterns (Church et al., 2010; Gastgeb, Dundas, Minshew, & Strauss, 2012), and drawn imaginary animals (Klinger & Dawson, 2001). These results suggest that participants with ASD may not use the information about similarity between objects during classification tasks. However, there are also reports of normal prototype abilities in ASD (Molesworth, Bowler, & Hampton, 2005), moreover, participants showing good and impaired performances in categorization tasks have been described within single studies (Molesworth, Bowler, & Hampton, 2008; Vladusich, Olu-Lafe, Kim, Tager-Flusberg, & Grossberg, 2010).

Measures of lexical development in ASD have revealed a number of atypicalities (Arunachalam & Luyster, 2016). Howlin (2003) contrasted patients diagnosed with autism and Asperger syndrome and found that the average age for production of first words was 38 months for the autism group and 15 months for the Asperger group, meanwhile first words in typically developing (TD) children usually appear around 12 months (Bates et al., 1988; Fenson et al., 1994; Tager-Flusberg et al., 2009). Both expressive and receptive language abilities appear delayed in populations with ASD, as compared with their TD peers, with some studies showing comprehension relatively more impaired than production (Hudry et al., 2010; but see the meta-analysis of Kwok, Brown, Smyth, & Oram Cardy, 2015, showing comparable delays in both domains).

A number of useful mechanisms for word learning have been characterized for the ASD population. For example, mutual exclusivity, known as the ability to map new words to previously unseen referents, is well preserved in ASD (de Marchena, Eigsti, Worek, Ono, & Snedeker, 2011; Preissler & Carey, 2005), as is the noun bias, a preference for mapping novel words to single objects instead of actions (Swensen, Kelley, Fein, & Naigles, 2007; Tek et al., 2008). Field, Allen, and Lewis (2016) observed that children with ASD and TD children prefer to extend novel words to objects that share the same function, and this function bias was observed when function competed with shape for label extension in both populations.

Studies have noted an absence of the shape bias in autism. Tek and colleagues (2008) used the intermodal preferential looking paradigm with a group of children with ASD with an average age of 33 months at the beginning of the study, and a group of TD children matched by vocabulary size. They assigned a novel word to a novel target object and later presented two novel objects to participants, one matching the shape but not the color of the target, and another matching its color but not its shape. Across four sets of evaluations, separated in time by 4 months each, there was a growing preference for looking at the shape match over the color match objects in the TD group, but no clear evidence of this bias in the ASD group. The authors noted that the absence of the shape bias in ASD was not due to small vocabulary size or to difficulties with the intermodal preferential looking paradigm and suggested that the age of appearance of the shape bias could be delayed in ASD. In an extension of this study, using a similar methodology, Potrzeba et al. (2015) evaluated the shape bias in a longitudinal study with a larger sample of children with ASD and a wider developmental time span: 20–54 months of age. They observed no shape bias in the overall performance of the ASD group but reported that some individuals in this group did show the bias.

Hartley and Allen (2014) evaluated generalization of words to new objects and object pictures in considerably older children with ASD: averaging 9.7 years of age. The children were tested in an object-sorting task: they were presented with a color picture of a new target object associated with a novel word (e.g. 'This is a blicket'). They were then asked to put all the blickets in a target box, and all the other objects in another box. They were presented with three different objects (and their corresponding pictures): one that matched the blicket in shape but not in color, another matching the blicket in color but not in shape, and a familiar object that did not share any of its physical properties. Four trials were presented to each participant. The results showed an absence of a shape bias in the ASD group: children with ASD used color as frequently as shape to sort the objects into boxes. In contrast, the control group of TD children, matched by verbal age, classified the objects based on their shape, without using color information. Neither TD children nor those with ASD extended new labels to familiar objects.

In all of these studies, the experimental stimuli to which participants with ASD extended new labels, ascertained either through looking times or their sorting of objects into different boxes, were all novel stimuli. It is thus a logical possibility that both the shape match and the color match objects were categorized as targets not because shape and color were equally weighted, as supposed by Hartley and Allen (2014), but because the target, the shape match, and the color match all shared the property of being novel stimuli.

In Experiment 1, we address this possibility by systematically replicating the object-sorting task of Hartley and Allen (2014), but with the inclusion of an extra novel object in each trial. This methodological variation allows us to evaluate two different hypotheses for the lack of the shape bias in autism. If children sort the color match and shape match objects into the target box, but not the novel object, that would support the hypothesis that shape and color information are equally important for label extension in ASD. In contrast, if the shape match, color match, and novel objects are all put in the target box, that would suggest an atypical categorization process where not only the physical similarity between objects, but also attention to novelty underlies the deficient label extension of children with ASD in the object-sorting task. One additional variation in our study is that we evaluated Spanish-speaking children. Studies on the shape bias have mainly assessed English-speaking children, but there is empirical evidence of attention to shape during label extension tasks in TD Spanish-speaking children (Colunga, Smith, & Gasser, 2009; Gathercole & Min, 1997).

In Experiment 2, we explore the object-sorting task in a computational model. Our objective is to present a biologically sound computational model of ASD to study the development and interaction of categorization and word-learning abilities. With this model, we attempt to understand the lack of a shape bias in ASD.

2 | EXPERIMENT 1

2.1 | Method

2.1.1 | Participants

Twenty-seven children with ASD (3 females) and 34 typically developing children (13 females) were evaluated in Experiment 1. Participants with ASD were recruited from two special education centers for children with ASD in Mexico City and the State of Mexico. All of them had been diagnosed with ASD by an expert clinician. Participants with ASD diagnosed with other neurological conditions were excluded from the sample, and we confirmed that participants with ASD had normal or corrected to normal vision and normal hearing. The TD children were recruited from a school in Mexico City, and from a database of parents who had expressed interest in participating in language studies of the Psycholinguistics Lab at the National Autonomous University of Mexico. All children were monolingual native Spanish speakers. TD children whose estimated verbal mental age (VMA) matched the VMA of children in the ASD group were included, as measured by the receptive vocabulary scale of the Peabody Picture Vocabulary Test-III (PPVT-III). None of the TD children had visual, hearing, or neurological problems according to parental reports. Informed consent was obtained from the parents of all participants. The Ethics Committee of the School of Psychology at the National Autonomous University of Mexico approved the study.

2.1.2 | Materials

Fifteen familiar objects were used for the training trials, and 16 unfamiliar objects and four familiar objects for the test trials (see Figure 1). All unfamiliar objects were created from pieces of wood, foam clay, or modeling clay. All objects were solid and hand-sized.

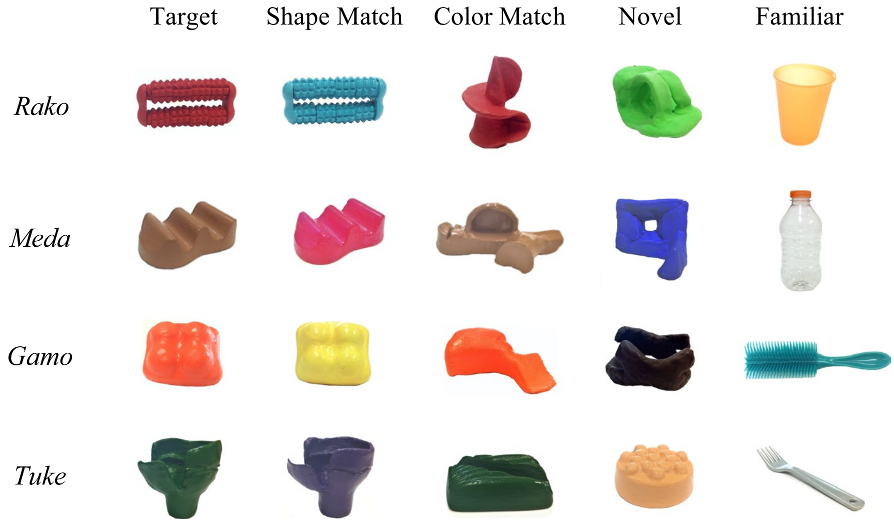

FIGURE 1 Experimental objects

Each of the four test trials included five objects: the target object, one shape match with the same shape as the target but a different color; one color match with the same color as the target but a different shape; one novel object: that was novel as the target, but with different shape and color; and one familiar object frequently used by children. Familiar objects in each trial did not share any properties with the target. In a previous pilot study run with six different TD children, none of them recognized the novel experimental objects as known objects. Participants classified the objects into one of two transparent boxes referred to as the target box and the non-target box.

Four new word labels, 'rako,' 'meda,' 'gamo,' and 'tuke,' were each assigned to the target objects, as shown in Figure 1. The word labels were created for this experiment with a two-syllable consonant-vowel structure (cvcv) that is very common in Spanish (Quilis, 1993). To capture the greater frequency of masculine gender nouns in Spanish, three of the four labels were marked with a masculine article preceding the noun (e.g. 'el tuke') and one with a feminine article ('la meda'). To prevent children from classifying the familiar objects into the non-target box because of gender mismatch, familiar objects were assigned to each trial to match the gender of the novel word label.

2.1.3 | Procedure

Our procedure was based on the study of Hartley and Allen (2014). Participants were individually tested, either in their schools or in a testing room at the School of Psychology at the National Autonomous University of Mexico. Each child sat at a table opposite the experimenter, with the two boxes on the table. After the experimenter introduced herself and established rapport, the next verbal instructions were delivered (in Spanish): "Today we are going to play with different objects that we need to sort into these two boxes." Two training trials were then administered.

Training trials

One of the two training trials used three toy cars as the target objects, along with three other hand-sized toys as distracters: a ball, a stuffed animal, and a plastic frog. The other training trial used four whiteboard markers as the target objects and three other stationary objects as distracters: an eraser, a protractor, and a children's pencil sharpener. The difference in the number of targets for each training trial was to prevent children from learning that equal numbers of objects should be classified into each box. The order of presentation of the two training trials was counterbalanced. The experimenter provided the following instructions: "Look at this object. It's a car (or a whiteboard marker in the other training trial). Can you see it?" The experimenter then handed the object to the child so she or he could manipulate it and continued with the instructions: "All cars go in this box [pointing to the target box]. I want you to put all cars in this box, and all other objects in the other box [pointing to the non-target box]. So, this car goes in this box. Can you put the car in this box?" After the child put the car in the target box the experimenter said: "Well done! Remember all cars go in this box, and all objects that are not cars go in this other box." The next object presented to the participant was a non-target object. The experimenter then asked the child "So, where does this object go?" The remaining objects were presented one by one in random order. After completing the first trial, all the objects were removed from the boxes and the second training trial started. The experimenter repeated all the instructions for the second set of objects.

Participants were praised for correct classification responses for each object during the training trials. The experimenter also reinforced good behavior and attention to instructions. Every time the participant made a mistake in classifying objects, the experimenter stopped the training trial, explained why the classification was wrong, the objects were removed from the boxes and the training trial started again. Training trials were stopped and repeated as many times as required. Only those participants completing the second training trial in no more than two attempts continued to the test phase.

Test trials

The four test trials were counterbalanced for order, and each trial followed the same structure used for the training trials, but without reinforcement for correct classifications.

Each test trial started with the presentation of the target object: the experimenter showed one of the target objects (e.g. a rako) to the participant and said, "This is a rako. See, it is a rako. Look carefully at the rako." Then, the experimenter pointed to the target box and explained: "Rakos go in this box. Since this one is a rako, it goes in this box. Can you put the rako in this box?" Then the experimenter said: "I want you to put all the objects that are not rakos in the other box." For each trial, the experimenter labeled the target object seven times. After the classification of the target object, the experimenter presented the remaining four objects of the trial, including one shape match, one color match, one novel object and one familiar object (Figure 1); these were presented one by one in random order. For these objects, the experimenter said to the participant: "Look at this one—where do you think this one goes? Can you put it in the correct box?" Children were considered to have extended the new word label to those objects classified into the target box.

The experimenter was trained in the application of the behavioral tasks, had previous experience working with TD and atypically developing children, and avoided behaviors that could interact with the children's responses.

2.2 | Results

2.2.1 | Training trials

Three boys with ASD did not continue to the test trials because they required more than two attempts to complete the second training trial. All TD children passed the two trials of the training phase. The final sample of participants in the ASD group included 21 boys and three girls. Consequently, we included the 21 boys and only three of

TABLE 1 Chronological age and estimated verbal mental age of participants in the TD and ASD groups

	TD (*n* = 24)			ASD (*n* = 24)			
	M	*SD*	Range	M	*SD*	Range	*p*-value
Age (years)	4.49	0.54	3.5 – 5.41	8.29	2.46	4.16 – 13.75	<0.001
Verbal mental age (years)	4.64	0.79	3.0 – 5.91	4.93	1.81	2.16 – 8.33	0.46

Abbreviations: ASD, autism spectrum disorders; TD, typically developing.

the 13 girls in the final sample of TD children. To do this, we selected the three TD girls who best matched the VMA of the three girls in the ASD group, the other 10 TD girls were excluded from the subsequent analyses.

Table 1 summarizes the characteristics of participants in both groups. A Student's *t*-test revealed no significant difference between groups in their verbal mental age (VMA; $p = 0.46$). For each group of participants, we analyzed correlations between the number of training trials and performance during tests. The only correlation approaching statistical significance was a positive one ($r = 0.401$, $p = 0.052$) between the number of training trials and the number of novel objects classified as targets, but this was observed only in the TD children, who rarely classified the novel objects as targets ($M = 0.16$).

2.2.2 | Test Trials

We calculated the number of times participants classified each type of object (target, shape match, color match, novel, or familiar) into the target box. The mean values for each group are shown in Figure 2.

A 2 (Group) × 4 (Object) mixed ANOVA was conducted. This analysis did not include measures of responses to the target objects because in all trials, children in both groups correctly classified all the target objects into the target box, so there was no variance in this condition. The analysis revealed significant main effects of Group $F(1, 46) = 13.35$, $p < 0.001$, Object $F(3, 138) = 191.16$, $p < 0.001$, and the interaction Group × Object $F(3, 138) = 8.17$, $p < 0.001$. A post-hoc analysis with the Bonferroni method was run to analyze pair-wise differences between and within groups. Children in the ASD and TD groups classified the shape match object into the target box at comparable rates (ASD $M = 3.79$, TD $M = 4$, $p = 1.0$), and significantly more often than any other object (all $ps < 0.001$). Children with ASD classified the color match objects into the target box ($M = 1.83$) significantly more often than their TD counterparts ($M = 0.54$, $p < 0.001$). These results are in accordance with those described by Hartley and Allen (2014). We did not observe differences in classifying the familiar objects between children with ASD ($M = 0.41$) and TD children ($M = 0$, $p = 1.0$) in the post-hoc analysis. Remarkably, in line with the hypothesis of children with autism paying attention to novelty, we found that children in the ASD group classified the novel objects into the target box significantly more often ($M = 1.25$) than the TD children ($M = 0.16$), who did so hardly at all ($p = 0.002$).[1]

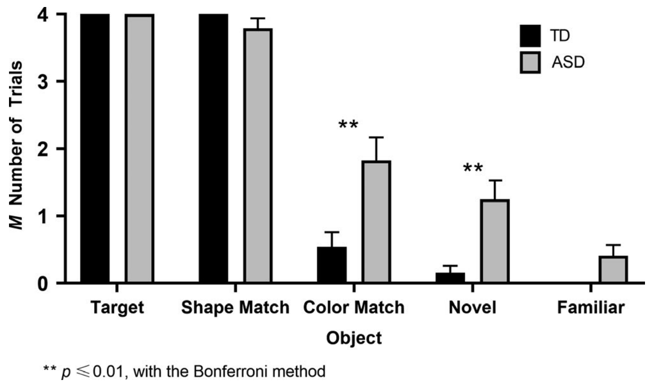

FIGURE 2 Mean number of trials for each object type that children with autism spectrum disorders and typically developing children classified into the target box. Error bars show the standard error of the mean

2.2.3 | Variability in ASD

The classification of objects in the ASD group was variable: some children showed a shape bias while others did not. To explore this variability and analyze whether errors in extending labels to color matches were systematically associated with label extensions to novel objects, we formed two subgroups of children with ASD: the Shape Bias subgroup, consisting of children who classified at least three shape match objects and no more than one color match object into the target box, and the Shape and Color subgroup, including participants who classified at least three shape match and three color match objects into the target box, as expected from the evidence of previous studies (Hartley & Allen, 2014; Tek et al., 2008). There were three outlier children who did not fall into either subgroup: one classified most of the objects into the non-target box, and the other two showed random responses across types of objects.

To evaluate the appropriateness of this grouping and prevent an experimenter bias in the data analysis, we conducted an unsupervised cluster analysis of the participants with ASD, based on their pattern of responses without a priori criteria. This analysis confirmed the existence of two main clusters highly convergent with our classification of Shape Bias and Shape and Color subgroups, and three outliers. There was a discrepancy, however, in the classification of two participants. One member of the Shape and Color subgroup was an outlier in the unsupervised clustering;

this participant classified all of the objects into the target box except for one familiar object. We took this participant out of the final Shape and Color subgroup. The cluster analysis also included another participant, who we had evaluated as responding randomly, in the Shape and Color subgroup. Since this participant only extended labels to two color matches, we were unable to include him in the Shape and Color subgroup. The final subgroups included 10 children each.

Participants in the Shape and Color subgroup consistently extended the new labels to both the shape and color matches in at least three of the four test trials. The mean number of classifications of each type of object for these subgroups is shown in Figure 3.

For each object type, the mean number of objects classified as targets between subgroups was compared with a Student's *t*-test. We were particularly interested in how the novel objects were treated by participants who used color information for label extension versus those who did not. Children in the Shape and Color subgroup treated the novel object as a target significantly more often ($M = 1.90$) than children in the Shape Bias subgroup ($M = 0.2$, $t(18) = -4.2$, $p < 0.001$). There were no significant differences between subgroups either in chronological age or VMA (Table 2).

2.3 | Discussion

Children with ASD were more likely than TD children to extend new word labels to objects based on their color. This result is consistent with previous descriptions of deficiencies in the shape bias for individuals with autism (Hartley & Allen, 2014; Potrzeba et al., 2015; Tek et al., 2008; Tek & Naigles, 2017). Notably, our study extends previous results in finding that the atypical label extension of children with ASD was more likely to reach novel objects that did not share the shape or color of the target object but were nonetheless novel as the target object.

We anticipated two possible scenarios of atypical label extension: in the first scenario, children with ASD would classify objects using the perceptual information of their shape and color; in the second, they would group the shape match, color match, and novel objects together.

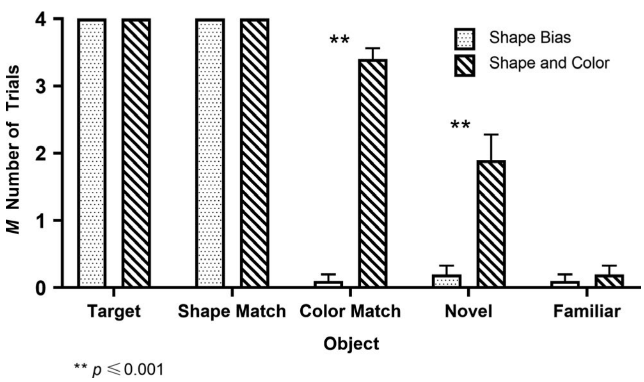

FIGURE 3 Mean number of trials for each object type that children with autism spectrum disorders in the shape bias and the shape and color subgroups classified into the target box. Error bars show the standard error of the mean

We found an intermediate scenario. Those children atypically using color information for label extension in at least three trials were more likely to classify the novel objects as targets, but the proportion of novel objects classified as targets was less than that of color matches. This result suggests that both factors – deficits in attention to the right physical properties of objects and increased attention to the novelty of objects – may contribute to the lack of a shape bias in autism, but neither factor alone fully explains the behavior. In Experiment 2, we explore the extent to which an atypical development of categorization processes in ASD accounts for deficits in the shape bias.

The high variability of responses within the ASD group limits the generalization of these results to the population with autism. However, the shape bias in our study appears as a reflection of the behavioral heterogeneity in ASD: we observed children with ASD categorizing objects with a shape bias like their TD peers, and others extending categories to objects that shared other irrelevant properties, such as color and novelty. On the one hand, this finding replicates the marked variability among individuals with ASD reported in different behavioral domains, including categorization (Molesworth et al., 2008), language development (Tek, Mesite, Fein, & Naigles, 2014), and the shape bias (Potrzeba et al., 2015). On the other hand, it raises the question of what other factors contribute to explain the different patterns of performance of children with ASD. Cimpian and Markman (2005) have suggested that the emergence of the shape bias in experimental studies is directly affected by methodological manipulations (e.g. object complexity, context); particularly, they have argued that the shape bias may be a task-induced artifact, this suggestion has been debated by Smith and colleagues (Kucker et al., 2019; Smith & Samuelson, 2006) who have argued that the way experiments are done and the effects of different methods on the participants' performance, are actually informative about the nature and development of the shape bias. Thus, future studies should evaluate the shape bias in ASD through different experimental tasks, analyze the atypical label extension to novel objects, and assess the extent to which methodological manipulations have differential effects in ASD. This would provide us with a more specific understanding of ASD and a comprehensive view of the shape bias in typical and atypical development.

Finally, there has been an emphasis on the relationship between development of the shape bias and vocabulary growth (Hahn & Cantrell, 2012; Smith et al., 2002). In our study, the TD and ASD groups were matched on vocabulary measures to ensure that differences between groups in the shape bias were not driven by differences in vocabulary knowledge. We matched participants on receptive vocabulary and found a discrepancy between chronological and verbal ages in ASD, with verbal age being delayed. This profile has been described before for children with ASD using measures of receptive vocabulary (Hartley & Allen, 2014) and both receptive and productive vocabularies (Tek et al., 2008). The slower acquisition of vocabulary in individuals with ASD who fail to show a shape bias suggests, as has been noted before (Tek et al., 2008), that children with ASD may rely on different (atypical) mechanisms for word learning.

TABLE 2 Chronological age and estimated verbal mental age of participants with ASD in the shape bias and the shape and color subgroups

	Shape bias subgroup (n = 10)			Shape and color subgroup (n = 10)			
	M	*SD*	Range	*M*	*SD*	Range	*p*-value
Age (years)	8.88	2.63	5.91 – 13.75	7.82	1.73	5.66 – 11.25	0.299
Verbal mental age (years)	5.81	2.15	2.16 – 8.33	4.43	1.21	3.16 – 6.91	0.094

Abbreviation: ASD, autism spectrum disorders.

3 | EXPERIMENT 2

In Experiment 2, we simulated the object-sorting task in a computational model (CM) running simulations of ASD and TD. Since the shape bias is at the intersection of visual categorization and word learning, we developed an artificial neural network model suitable for forming categories of visual stimuli and learning mappings between categories of objects and word labels. A key component of our neural model is a Self-Organizing Map (SOM), consisting of a grid of artificial neurons (processing units) used to represent and categorize visual stimuli in different clusters (Kohonen, 2013).

Our CM was grounded in previous work modeling categorization and lexical development, in particular that of Mayor and Plunkett (2010) using two SOMs, one of which processed visual stimuli (e.g. pictures of dogs), and the other auditory stimuli (e.g. the word 'dog'). These maps were connected with Hebbian connections, so that frequent coactivations were reinforced between units in the visual and auditory maps. The model of Mayor and Plunkett was trained, first by presenting the corresponding maps with different visual and auditory stimuli, and then by presenting a large number of picture-word pairings. Some phenomena were observed in the functioning of this model that resembled word-learning processes. For example, presentation to the auditory map of the word 'dog' produced spreading activation through the Hebbian connections that activated the units representing a prototypical dog on the visual map, simulating word comprehension. With this model, word-learning processes were simulated, including fast mapping, taxonomic responding, and the initial advantage of comprehension over production of words.

3.1 | Overview of the computational simulations

Our computational model is an abstract one, intended to explore the development of representations of visual objects and categories, their organization on a map of representations, and their interaction with verbal labels. We modeled ASD by simulating atypical neuronal connectivity on the SOM, and we analyzed the impact of these low-level atypical mechanisms on the high-level complex cognitive processes of categorization and emergence of the shape bias. Details of the model implementation are presented in Appendix A.

Before being exposed to the object-sorting task, the CMs went through a preexperimental phase, where they were exposed to input vectors representing 88 visual objects belonging to one of 30 frequently used categories (e.g. bottles, crayons, telephones), and also to object-word pairings. This was done to provide the models with developmental trajectories in categorization and lexical acquisition, to mirror children's experience before they are exposed to the experimental tasks.

After the preexperimental phase, the 20 objects used in the object-sorting task of Experiment 1 were presented to the CMs, and performance of the simulations was analyzed to determine whether the shape match, color match, novel, and familiar objects were categorized in the same cluster with the target object in the SOM, or outside this cluster.

3.1.1 | Model architecture

Our model is composed of one SOM dedicated to forming and organizing categories of visual stimuli, and one layer of units representing word labels. The SOM and the word-label layer communicate through Hebbian connections (Figure 4).

3.1.2 | SOM functioning

The SOM is composed of 225 units, arranged in a 15 x 15 grid of processing units (Figure 4, right). Each unit is characterized by a weight vector that allows for the processing of stimuli, and the adaptation of weight vectors captures learning in the SOM. During the preexperimental phase, the units in the SOM become attuned to processing certain input patterns that represent visual objects. With this training, a topological structure is expected to emerge in the SOM, so that units processing similar objects will be closer to one another on the map, while units processing dissimilar objects will be more distant. In this way, the proximity of objects represented on the map can be used to infer how they are categorized; those objects processed in nearby units are considered members of the same category.

The topological organization of the map results from the dynamics of *competition* and *cooperation* between processing units. When a visual object is presented as an input to our model (e.g. a cup), the units compete, and the unit best adapted to process cups becomes the winning neuron. The location of this unit on the map represents the place where cups are better processed. Then cooperation starts: a neighborhood of units is established around the winning unit (Figure 4), and those units within the neighborhood will adapt their weight vectors to become better at processing cups. After many presentations of different inputs, the map becomes topologically organized.

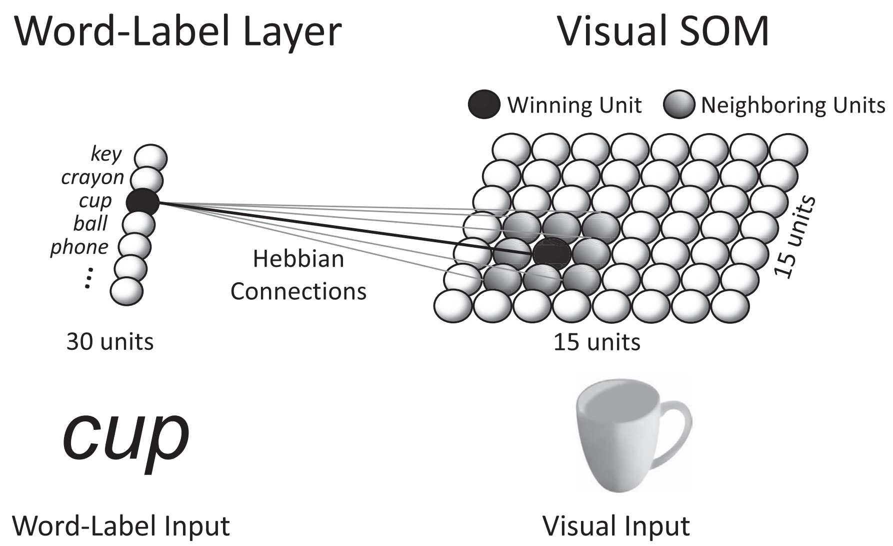

FIGURE 4 Architecture of the computational model used in Experiment 2. Only some Hebbian connections are shown for clarity

3.1.3 | Modeling TD and ASD in the SOM

Typically, the learning rate (α) of units and the size of the cooperation neighborhood (*N*) in SOMs are initially large and decay over time (Kohonen, 2013). In our model of TD, we used standard functions for α and *N* (see Appendix A). Our modeling of ASD consisted of favoring local hyperplasticity and hyperconnectivity in the SOM. We captured local hyperplasticity with a non-decaying learning rate α that guaranteed a continuous strong adjustment of weight vectors in the SOM, and local hyperconnectivity was modeled with a non-decaying *N* function with stronger cooperation for closer neighboring units and weaker cooperation for more distant ones than in the TD simulation (Figure A1).

3.1.4 | Biological motivation of the ASD model

Our computational modeling of ASD, while abstract and simple, reflects numerous neurological dysfunctions that have been put forward to explain autism. Local hyperplasticity has been observed in animal models of ASD; when compared with controls, they show a significantly superior synaptic efficiency in different brain regions, including somatosensory cortex, prefrontal cortex, and amygdala (Markram & Markram, 2010; Rinaldi, Perrodin, & Markram, 2008). Local hyperconnectivity has been directly observed in cortical areas of mouse models of ASD (Nagode et al., 2017; Silva et al., 2009); and has been suggested as leading to exaggerated recruitment of neighboring units, probably accounting for the hyperreactivity in local neuronal circuits found in autism (Markram et al., 2007). These patterns of connectivity cannot be directly measured in human brains, but recent analyses of structural covariance networks performed in a large dataset of brain images of individuals with autism show a balance strongly favoring local short-range over global long-range connectivity (Bethlehem, Romero-Garcia, Mak, Bullmore, & Baron-Cohen, 2017).

Our manipulation of both α and *N* leads to a hyperexcitable state of the SOM, in line with one of the most widely accepted neuronal theories of autism, which postulates an increased excitation/inhibition ratio in neural systems as the basis of many of the symptoms observed in this condition (Rubenstein & Merzenich, 2003).

3.1.5 | Word-label layer and object-word pairings

Words are simulated with binary patterns in one layer of 30 symbolic units (Figure 4, left); the activation of one unit simulates the auditory presentation of the word the unit represents.

The word-label layer is fully connected to the SOM. When a visual object is presented to the SOM concurrently with the activation of a unit in the word-label layer, a strong coactivation is detected and the connection between highly coactive units strengthens, following Hebbian learning principles (see Appendix A).

Once the Hebbian connections are strengthened (i.e. they acquire a value >0.8 in the range 0–1), they allow spreading activation from the word-label layer to the SOM. This process permits the SOM organization to be modulated by both word labels and visual inputs. During the presentation of the object-word pairing of a cup, for example, the activation of the word ‘cup’ propagates through strong Hebbian connections to activate the last stored representation of cups in the SOM. The activated units in the SOM are then used to process the current visual input of the cup, and the representation of cups is also updated in the SOM with the presented exemplar. The effect of a label in the model consists of recovering the strongest previous representation of the corresponding visual object, which is used to process the visual input more efficiently.

Our implementation of the Hebbian algorithm is based on a biologically motivated model of associative learning presented by Tovar, Westermann and Torres (2018). The parameters used for our simulations are indicated in Appendix A. To keep our simulations simple,

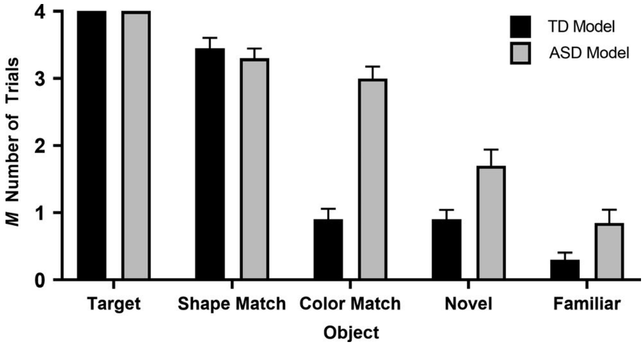

FIGURE 5 Mean number of trials classified into the target box for each object type in the computational simulations of autism spectrum disorders and typically developing children. Error bars show the standard error of the mean

and since our main goal is to analyze the role of atypical categorization of visual objects in ASD, we used the same parameters for the Hebbian connections in the TD and ASD models.

3.1.6 | Stimuli

The preexperimental phase used 88 objects and the test phase 20 objects. The four familiar objects used in the test phase were also used in the preexperimental phase. Objects were coded as 7-dimensional vectors composed of four shape dimensions (length, width, height, number of features), and three color dimensions (red, green, blue). Details of the coding of objects are presented in Appendix A.

The 88 objects in the preexperimental phase belonged to one of 30 categories corresponding to frequently used objects. Each category had from two to five exemplars. For most categories, there were minimal variations in the shape of different exemplars, but color was more variable: crayons, for example, were all about the same size, but one was purple, one was yellow, and one was orange. Word labels corresponded to the basic level category names (e.g. bottle, crayon, telephone); all crayons, for example, were labeled with the same word.

3.1.7 | Preexperimental phase

This phase was divided into two stages. Based on the work of Mayor and Plunkett (2010), our model included unsupervised and supervised learning. In the first stage, the visual SOM formed categories based on the similarities between the preexperimental objects. This procedure captures the child's perceptual refinement that results mainly from unsupervised learning activities, as it is the direct exposure to objects – and consisted of 400 epochs, each epoch included one presentation of all the preexperimental objects.

In the second stage, the organization of the SOM was modulated by both the visual objects and their word labels as explained above; this stage constitutes the supervised learning component of our simulations, as it captures social activities where children are exposed to objects and the labels that others use to refer these objects. This stage consisted of 100 epochs. One important difference between our simulations and those of Mayor and Plunkett (2010) is that in their model, word labels did not affect the processing of categories in the visual SOM. In our simulations, words modulated visual representations in the last 100 epochs. This modulatory effect has been suggested as a more realistic mechanism for the development of categorization and stimulus representations, as shown by the computational simulations of Westermann and Mareschal (2014).

3.1.8 | Experimental phase

The input patterns representing the five objects of each of the four testing trials of Experiment 1 were presented. The performance of the model in this phase was evaluated in terms of how the experimental objects were processed in the SOM. Objects processed in neighboring units with the target (within a radius ≤5) were considered to be categorized together. Objects processed by distant units in the SOM were considered to belong to different categories. Twenty simulations of TD and 20 of ASD were performed.

3.2 | Results and discussion

Our computational simulations replicated the pattern of group results from Experiment 1. Most of the TD simulations showed a shape bias, while most of the ASD simulations did not. Figure 5 shows the mean number of categorizations of each type of object as a target across the 20 runs of the TD and ASD models. Using the same criteria as Experiment 1, we categorized our computational simulations into a Shape Bias subgroup and a Shape and Color subgroup. Sixteen of the 20 runs of the TD model showed a shape bias, and 12 of the 20 in the ASD model met the shape and color criterion; the remaining runs did not meet the criteria for either subgroup. We then compared these subgroups.

We analyzed the map locations for processing the experimental objects. In the TD simulations, the target and shape match objects were processed in nearby locations and the color match and novel objects were usually more distant, generating a more distributed processing of experimental objects in the maps (Figure 6, right column). In the ASD simulations, the experimental objects were frequently classified by nearby units, showing that they were all treated as similar objects.

3.2.1 | Organization of categories in the model

To evaluate whether deficits in the development of categorization abilities underlie atypicalities in the shape bias in the ASD model, we analyzed the developmental trajectory of preexperimental categories in the visual maps. This was done for two developing stages: one after 400 epochs, when the maps had considerable input to refine visual categorization but no experience with object-word pairings, and a second one at the end of the preexperimental phase, after 400 epochs of visual input plus 100 epochs of object-word pairings. Examples of visual representations of the categories formed by the TD and ASD models are presented in Figure 6.

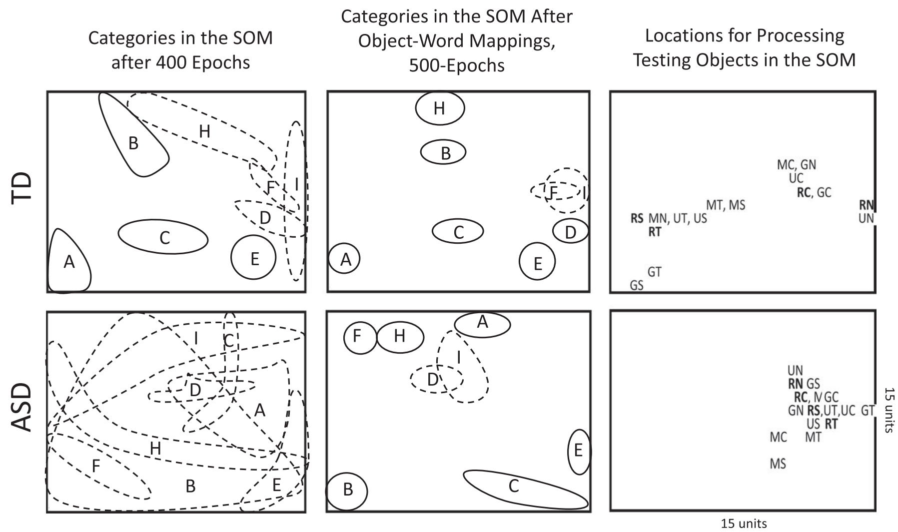

FIGURE 6 Six square representations of self-organizing maps (SOMs). Columns 1 and 2 show example locations where preexperimental categories were processed after 400 epochs (column 1) and after 500 epochs (column 2). To simplify the figure, only eight categories are shown, indicated by letters A-I. Lines in the SOMs delimit the areas where objects in the categories were processed. Continuous lines show well-delimited categories; dashed lines show overlapping categories. Column 3 shows example locations for processing the testing objects. Each testing object is indicated by two letters: the first indicates the trial (R = rako, *M* = meda, G = gamo, U = tuke), and the second indicates the type of object (T = target, S = shape match, C = color match, *N* = novel object). The objects of the rako trial are in bold to highlight the higher clustering in the ASD simulation compared with the TD simulation

After 400 epochs, categories show a developing organization in the TD model: objects belonging to the same categories are grouped closer on the map, and there is minimal overlap between categories. In sharp contrast, the organization of categories in the ASD model is deficient: objects from the same categories are processed in distant parts of the map, revealing fragmented representations of categories, and there are many overlapping categories. This organization was expected, based on the theoretical work of Gustafsson (1997); moreover, we provide computational evidence for his theoretical assumption that simulating autism with SOMs may result in a fragmented organization of representations on the maps.

A second outcome, expected particularly in the TD simulation as a result of previous modeling work by Westermann and Mareschal (2014), was the observation of more compressed representations of categories after object-word pairings. As shown in Figure 6, the learning of object-word mappings favored the delimitation of categories. To eliminate the possibility that this result was just an effect of more exposure to visual input, we ran an additional set of simulations for the same 500 epochs without presenting labels in the last 100, and observed sparser final representations of categories. Compressed representations of categories emerged as well in the ASD simulation after object-word pairings. Presentation of a label started a competition between the units in the SOM representing the named category, the winning representation attracted the processing of incoming visual objects. In the TD maps with competing representations of a category that were already close (organized), this process favored a refinement of the categorical representation. In the ASD maps, however, competition between fragmented representations of the named category produced one of the fragments that attracted the processing of the incoming objects of the category, preventing a rich integration of previous representations of the category exemplars. One result of this poor integration of exemplars is an impaired formation of prototypical representations of categories, which is consistent with evidence from empirical studies of autism (Church et al., 2010; Gastgeb et al., 2012, 2009; Klinger & Dawson, 2001). Notably, our model provides insight into mechanisms for the impaired formation of prototypes in ASD.

Our computational simulations are also congruent with two empirical findings that have been difficult to integrate, one establishing a strong relation between the use of the shape bias as a mechanism for word learning (Smith et al., 2002), and the other showing that children with autism can develop sizeable lexicons even when they do not show a shape bias (Tek et al., 2008), as we also observed in Experiment 1. Our model suggests that word-object mappings can be established when the associative learning abilities are well preserved in ASD, even when categories are not well organized around shape properties. Moreover, the model sheds light on how with a lexical development like this one, the representations of word referents in autism may be idiosyncratic; for example, referents may be highly specific, when the word label is strongly associated with one particular instance (fragment) of a category of objects, or quite general and inaccurate, when the word label is associated with a poor (fragmentary) representation of the category.

3.2.2 | Processing shape and color

There are two factors accounting for the shape bias in our TD simulation. The first concerns the statistical properties of input vectors. Inputs consisted of seven dimensions, with more dimensions (four) coding for shape than those (three) coding for color. Consequently, when shape and color compete during the topological organization of the SOM, learning based on shape results in a greater reduction in error between the current input and the current representation of this object. The adaptation of the SOM based primarily on shape is thus more efficient, and organization of color dimensions is subordinated to shape. The result is a topological organization of shape and a distributed organization of color in the SOM. Units on the upper left of the SOM, for example, could be optimally attuned to processing large objects and those on the lower right to processing small objects, but with units moderately attuned on both sides to process different colors. We argue that the statistical distinction between shape and color in our input vectors qualitatively captures the physical variations of shape and color of real objects; shape is a construct made up of many more properties than color.

The second factor is that word labels boosted the organization of the SOM based on shape. For most categories, objects named with the same word had minimal variations in shape but greater variations in color. The use of words in the simulations recovered previous representations of similarly shaped objects for the processing of incoming inputs, grouping representations of category members around shape and lessening the influence of color.

The continuous noise produced by the hyperplasticity and hyperconnectivity in the ASD model impaired a fine topological organization of shape features in the SOM. An adequate topological organization of shape dimensions, like that emerging in the TD simulations, provides the model with an efficient discrimination-categorization balance, distinguishing similar objects by processing them in distinct but nearby units. The atypical topological organization that emerges in the ASD models impairs this balance. The atypical organization in the Shape and Color subgroup of the ASD simulations may be an indication that the similarity between novel experimental objects was more important than their differences, making it more likely to categorize experimental objects together. This suggests that novelty of experimental objects may be an important factor in explaining atypical categorization in ASD.

The functioning of our model is congruent with the Attentional Learning Account (Colunga & Smith, 2008; Smith & Samuelson, 2006), in that it explains typical and atypical behavior based on domain-general mechanisms for learning. In particular, it is sensitive to the correlations among linguistic devices (labels), object properties (e.g. shape), and the organization of perceptual categories (in the SOM). With these mechanisms, we were able to account for the shape bias observed in TD children and the lack of a shape bias in children with autism.

4 | GENERAL DISCUSSION

This study used an experiment with children and computational simulations to analyze the previously reported lack of a shape bias in English-speaking children with autism (Hartley & Allen, 2014; Potrzeba et al., 2015; Tek et al., 2008). We have replicated this evidence and extended it to Spanish-speaking children with autism. We observed that such children who extended word labels to objects based on color were also more likely to extend labels to novel objects that did not share shape or color properties with the target objects. Our empirical data suggest that in addition to using irrelevant information such as color for label extension, the lack of a shape bias in children with ASD may be explained, at least in part because of a tendency to categorize experimental objects based on their novelty. This atypical treatment of experimental objects could be driven by an underlying atypical development of categorization in ASD, as demonstrated by our computational simulations.

We argue that a functional network underlying the shape bias requires building mature visual and verbal representations, and efficient links between them. While our simulations detail one route from hyperconnectivity and hyperplasticity to disruptions in visual category representations, with consequent deficits in the shape bias, we do not conclude that all deficits in the shape bias result from these altered mechanisms. Within the framework of our model, composed of three main components – the visual SOM, the word-label layer, and the Hebbian connections – only the visual SOM was altered to simulate ASD. Specific lesions to the other components could potentially affect the shape bias. Two recent papers that reported a weakened shape bias (Collisson, Grela, Spaulding, Rueckl, & Magnuson, 2015) and a strong novelty bias (Kucker, McMurray, & Samuelson, 2018) have suggested underlying mechanisms in line with this approach; the latter observed a preference for novel objects during referent selection tasks in two experiments and a set of simulations of young TD children (18 months of age). This novelty bias may be comparable to the label extension to novel objects in children with ASD documented in our study: remarkably, the authors suggested that the novelty bias they observed may result from unrefined lexicons in poorly organized associative networks of words and referents during early stages of vocabulary development. It is worth noting that they analyzed vocabulary growth through productive vocabulary. Our study, however, used measures of receptive vocabulary. Future research should analyze the relationship between a novelty bias and productive vocabularies in ASD.

Collisson et al. (2015) evaluated children with specific language impairment (SLI) and found that they show a less robust shape bias than TD children, and that abilities in associative learning are a strong predictor of their performance. They suggest that weaker associative learning abilities may contribute to weakness in detecting the covariations between visual and linguistic inputs, affecting the shape bias. Future modeling work of SLI could test this hypothesis by simulating lesions to the Hebbian connections in our model.

The results of Experiments 1 and 2 provide some suggestions for clinical practice. Children with ASD show comparable vocabulary as their TD counterparts but at significantly older chronological ages, suggesting that language interventions in ASD based on training the shape bias could favor more rapid lexical acquisition. Smith et al. (2002) demonstrated in TD children that training the shape bias in laboratory conditions generalizes to objects outside the laboratory and significantly increases their lexical acquisition. In line with this evidence, our data suggest that future clinical interventions for ASD should test the effects of training programs to improve the shape bias by mapping word labels to novel objects of the same shape category. This training should also include the explicit rejection of nonmembers of the category, with objects that share similar physical properties with the target objects, such as color, that are irrelevant to label extension, and it may be critical to also train the rejection of objects that share non-physical properties, like novelty.

We have proposed a computational model to explore the interaction between atypical perceptual categorization and word learning, and with this model we put forward a comprehensive account of the lack of a shape bias in autism. Notably, our model summarizes and incorporates findings from altered neural connectivity in ASD, theoretical and computational accounts of word learning, and behavioral descriptions of ASD (Colunga & Smith, 2008; Gustafsson, 1997; Hartley & Allen, 2014; Markram & Markram, 2010; Mayor & Plunkett, 2010; Rubenstein & Merzenich, 2003; Westermann & Mareschal, 2014) into a coherent account of atypical categorization and shape bias in autism. In doing so, our model also strengthens a domain-general account to understand the shape bias in typically developing individuals as a complex process emerging from the interaction of perceptual categorization and word learning based on associative learning mechanisms.

We argue that the lack of a shape bias and delays in lexical development in ASD do not represent deficits in domain-specific abilities; instead, our model postulates a route from general low-level atypical mechanisms to complex cognitive processes, where hyperconnectivity and hyperplasticity of local networks in ASD impair an efficient grouping of perceived objects based on their similarities, that produces fragmented representations of categories, and impairs the efficient linking of category representations and word labels. Our simulations also describe a typical developmental trajectory that bridges categorization with word learning: first, the covariances of object properties are detected, and when more characteristics of objects code for shape than color, shape is more important to the incipient perceptual categorization. Correlations with word labels then boost the grouping of objects around those object properties that covariate with labels, favoring the emergence of the shape bias.

ACKNOWLEDGMENTS

This work was funded by a research grant to the first author from CONACYT, México (CB 285152), and by a DGAPA-PAPIIT-IN304417 to the third author. We thank the children and families who participated in the study, Caritas de Amistad, A.C, and Centro Autismo Teletón.

CONFLICT OF INTEREST

None of the authors have any interests that might be interpreted as influencing the research.

DATA AVAILABILITY STATEMENT

The data and computer codes that support the findings of this study are available from Ángel Eugenio Tovar upon request (aetovar@unam.mx).

ENDNOTE

[1] The two groups of participants were matched on VMA, but we nonetheless analyzed whether there was a correlation between VMA scores and their performance in the experimental task. All the correlations were non-significant and weak to moderate; however, for the group with ASD, the correlations suggested that as VMA scores increased the performance in the sorting task was better; participants with ASD showed a positive correlation between VMA and number of shape match objects classified as targets ($r = 0.359$, $p = 0.085$), and negative correlations between VMA and: color match objects classified as targets ($r = -0.313$, $p = 0.136$); novel objects classified as targets ($r = -0.057$, $p = 0.79$); and familiar objects classified as targets ($r = -0.355$, $p = 0.085$). The TD group showed a weak negative correlation between VMA and color match objects classified as targets ($r = -0.034$, $p = 0.873$); and a weak negative correlation between VMA and novel objects classified as targets ($r = -0.167$, $p = 0.434$). Correlations for this group between VMA and shape match and familiar objects were not analyzed because there was no variability in the classification of these objects across participants.

REFERENCES

Arunachalam, S., & Luyster, R. J. (2016). The integrity of lexical acquisition mechanisms in autism spectrum disorders: A research review. *Autism Research*, *9*(8), 810–828. https://doi.org/10.1002/aur.1590

Bates, E., Bretherton, I., Snyder, L., Beeghly, M., Shore, C., McNew, S., … Garrison, A. (1988). *From first words to grammar: Individual differences and dissociable mechanisms*. New York, NY: Cambridge University Press.

Bethlehem, R. A. I., Romero-Garcia, R., Mak, E., Bullmore, E. T., & Baron-Cohen, S. (2017). Structural Covariance Networks in Children with Autism or ADHD. *Cerebral Cortex*, *27*(8), 4267–4276. https://doi.org/10.1093/cercor/bhx135.

Church, B. A., Krauss, M. S., Lopata, C., Toomey, J. A., Thomeer, M. L., Coutinho, M. V., … Mercado, E. (2010). Atypical categorization in children with high-functioning autism spectrum disorder. *Psychonomic Bulletin & Review*, *17*(6), 862–868. https://doi.org/10.3758/PBR.17.6.862

Cimpian, A., & Markman, E. M. (2005). The Absence of a Shape Bias in Children's Word Learning. *Developmental Psychology*, *41*(6), 1003–1019. https://doi.org/10.1037/0012-1649.41.6.1003

Collisson, B. A., Grela, B., Spaulding, T., Rueckl, J. G., & Magnuson, J. S. (2015). Individual differences in the shape bias in preschool children with specific language impairment and typical language development: Theoretical and clinical implications. *Developmental Science*, *18*(3), 373–388. https://doi.org/10.1111/desc.12219

Colunga, E., & Smith, L. B. (2008). Knowledge embedded in process: The self-organization of skilled noun learning. *Developmental Science*, *11*(2), 195–203. https://doi.org/10.1111/j.1467-7687.2007.00665.x

Colunga, E., Smith, L. B., & Gasser, M. (2009). Correlation versus prediction in children's word learning: Cross-linguistic evidence and simulations. *Language and Cognition*, *1*(2), 197–217. https://doi.org/10.1515/LANGCOG.2009.010

de Marchena, A., Eigsti, I.-M., Worek, A., Ono, K. E., & Snedeker, J. (2011). Mutual exclusivity in autism spectrum disorders: Testing the pragmatic hypothesis. *Cognition*, *119*(1), 96–113. https://doi.org/10.1016/j.cognition.2010.12.011

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., … Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5), 174–185. https://doi.org/10.2307/1166093

Field, C., Allen, M. L., & Lewis, C. (2016). Are Children with Autism Spectrum Disorder Initially Attuned to Object Function

Rather Than Shape for Word Learning? *Journal of Autism and Developmental Disorders*, *46*, 1210–1219. https://doi.org/10.1007/s10803-015-2657-5

Gastgeb, H. Z., Dundas, E. M., Minshew, N. J., & Strauss, M. S. (2012). Category Formation in Autism: Can Individuals with Autism Form Categories and Prototypes of Dot Patterns? *Journal of Autism and Developmental Disorders*, *42*(8), 1694–1704. https://doi.org/10.1007/s10803-011-1411-x

Gastgeb, H. Z., Rump, K. M., Best, C. A., Minshew, N. J., & Strauss, M. S. (2009). Prototype formation in autism: Can individuals with autism abstract facial prototypes? *Autism Research*, *2*(5), 279–284. https://doi.org/10.1002/aur.93

Gathercole, V. C. M., & Min, H. (1997). Word meaning biases or language-specific effects? Evidence from English, Spanish and Korean. *First Language*, *17*(49), 31–56. https://doi.org/10.1177/014272379701705102

Gustafsson, L. (1997). Inadequate cortical feature maps: A neural circuit theory of autism. *Biological Psychiatry*, *42*(12), 1138–1147. https://doi.org/10.1016/S0006-3223(97)00141-8

Hahn, E. R., & Cantrell, L. (2012). The shape-bias in Spanish-speaking children and its relationship to vocabulary. *Journal of Child Language*, *39*(2), 443–455. https://doi.org/10.1017/S030500091100016X

Hartley, C., & Allen, M. L. (2014). Brief report: Generalisation of word-picture relations in children with autism and typically developing children. *Journal of Autism and Developmental Disorders*, *44*(8), 2064–2071. https://doi.org/10.1007/s10803-014-2074-1

Howlin, P. (2003). Outcome in high-functioning adults with autism with and without early language delays: Implications for the differentiation between autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *33*(1), 3–13.

Hu, V. W., & Steinberg, M. E. (2009). Novel clustering of items from the Autism Diagnostic Interview-Revised to define phenotypes within autism spectrum disorders. *Autism Research*, *2*(2), 67–77. https://doi.org/10.1002/aur.72

Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., & Aldred, C., … PACT Consortium (2010). Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. *International Journal of Language & Communication Disorders*, *45*(6), 681–690. https://doi.org/10.3109/13682820903461493

Hupp, J. M. (2015). Development of the shape bias during the second year. *The Journal of Genetic Psychology*, *176*(1–2), 82–92. https://doi.org/10.1080/00221325.2015.1006563

Klinger, L. G., & Dawson, G. (2001). Prototype formation in autism. *Development and Psychopathology*, *13*(1), 111–124. https://doi.org/10.1017/S0954579401001080

Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks*, *37*, 52–65. https://doi.org/10.1016/j.neunet.2012.09.018

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2018). Too Much of a Good Thing: How Novelty Biases and Vocabulary Influence Known and Novel Referent Selection in 18-Month-Old Children and Associative Learning Models. *Cognitive Science*, *42*(Suppl 2), 463–493. https://doi.org/10.1111/cogs.12610

Kucker, S. C., Samuelson, L. K., Perry, L. K., Yoshida, H., Colunga, E., Lorenz, M. G., & Smith, L. B. (2019). Reproducibility and a unifying explanation: Lessons from the shape bias. *Infant Behavior and Development*, *54*, 156–165. https://doi.org/10.1016/j.infbeh.2018.09.011

Kwok, E. Y. L., Brown, H. M., Smyth, R. E., & Oram Cardy, J. (2015). Meta-analysis of receptive and expressive language skills in autism spectrum disorder. *Research in Autism Spectrum Disorders*, *9*, 202–222. https://doi.org/10.1016/j.rasd.2014.10.008

Markram, H., Rinaldi, T., & Markram, K. (2007). The Intense World Syndrome – an Alternative Hypothesis for Autism. *Frontiers in Neuroscience*, *1*(1), 77–96. https://doi.org/10.3389/neuro.01.1.1.006.2007

Markram, K., & Markram, H. (2010). The intense world theory - a unifying theory of the neurobiology of autism. *Frontiers in Human Neuroscience*, *4*, 224. https://doi.org/10.3389/fnhum.2010.00224

Mayor, J., & Plunkett, K. (2010). A neurocomputational account of taxonomic responding and fast mapping in early word learning. *Psychological Review*, *117*(1), 1–31. https://doi.org/10.1037/a0018130

Molesworth, C. J., Bowler, D. M., & Hampton, J. A. (2005). The prototype effect in recognition memory: Intact in autism? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *46*(6), 661–672. https://doi.org/10.1111/j.1469-7610.2004.00383.x

Molesworth, C. J., Bowler, D. M., & Hampton, J. A. (2008). When Prototypes Are Not Best: Judgments Made by Children with Autism. *Journal of Autism and Developmental Disorders*, *38*(9), 1721–1730. https://doi.org/10.1007/s10803-008-0557-7

Nagode, D. A., Meng, X., Winkowski, D. E., Smith, E. D., Khan-Tareen, H., Kareddy, V., … Kanold, P. O. (2017). Abnormal development of the earliest cortical circuits in a mouse model of Autism Spectrum Disorder. *Cell Reports*, *18*(5), 1100–1108. https://doi.org/10.1016/j.celrep.2017.01.006

Potrzeba, E. R., Fein, D., & Naigles, L. (2015). Investigating the shape bias in typically developing children and children with autism spectrum disorders. *Frontiers in Psychology*, *6*, https://doi.org/10.3389/fpsyg.2015.00446

Preissler, M. A., & Carey, S. (2005). The role of inferences about referential intent in word learning: Evidence from autism. *Cognition*, *97*(1), B13–23. https://doi.org/10.1016/j.cognition.2005.01.008

Quilis, A. (1993). *Tratado de fonología y fonética españolas*. Madrid: Gredos.

Rinaldi, T., Perrodin, C., & Markram, H. (2008). Hyper-Connectivity and Hyper-Plasticity in the Medial Prefrontal Cortex in the Valproic Acid Animal Model of Autism. *Frontiers in Neural Circuits*, *2*, https://doi.org/10.3389/neuro.04.004.2008

Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: Increased ratio of excitation/inhibition in key neural systems. *Genes, Brain, and Behavior*, *2*(5), 255–267. https://doi.org/10.1034/j.1601-183X.2003.00037.x

Samuelson, L. K. (2002). Statistical regularities in vocabulary guide language acquisition in connectionist models and 15–20-month-olds. *Developmental Psychology*, *38*(6), 1016–1037. https://doi.org/10.1037/0012-1649.38.6.1016

Silva, G. T., Le Bé, J.-V., Riachi, I., Rinaldi, T., Markram, K., & Markram, H. (2009). Enhanced Long-Term Microcircuit Plasticity in the Valproic Acid Animal Model of Autism. *Frontiers in Synaptic Neuroscience*, *1*, https://doi.org/10.3389/neuro.19.001.2009

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, *13*(1), 13–19. https://doi.org/10.1111/1467-9280.00403

Smith, L. B., & Samuelson, L. (2006). An attentional learning account of the shape bias: Reply to Cimpian and Markman (2005) and Booth, Waxman, and Huang (2005). *Developmental Psychology*, *42*(6), 1339–1343. https://doi.org/10.1037/0012-1649.42.6.1339

Swensen, L. D., Kelley, E., Fein, D., & Naigles, L. R. (2007). Processes of language acquisition in children with autism: Evidence from preferential looking. *Child Development*, *78*(2), 542–557. https://doi.org/10.1111/j.1467-8624.2007.01022.x

Tager-Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., … Yoder, P. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research*, *52*(3), 643–652. https://doi.org/10.1044/1092-4388(2009/08-0136)

Tek, S., Jaffery, G., Fein, D., & Naigles, L. R. (2008). Do children with autism spectrum disorders show a shape bias in word learning? *Autism Research*, *1*(4), 208–222. https://doi.org/10.1002/aur.38

Tek, S., Mesite, L., Fein, D., & Naigles, L. (2014). Longitudinal analyses of expressive language development reveal two distinct language profiles among young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *44*(1), 75–89. https://doi.org/10.1007/s10803-013-1853-4

Tek, S., & Naigles, L. (2017). The shape bias as a word-learning principle: Lessons from and for autism spectrum disorder. *Translational Issues in Psychological Science*, *3*(1), 94–103. https://doi.org/10.1037/tps0000104

Tovar, Á. E., Westermann, G., & Torres, A. (2018). From altered synaptic plasticity to atypical learning: A computational model of Down syndrome. *Cognition*, *171*, 15–24. https://doi.org/10.1016/j.cognition.2017.10.021

Vladusich, T., Olu-Lafe, O., Kim, D.-S., Tager-Flusberg, H., & Grossberg, S. (2010). Prototypical category learning in high-functioning autism. *Autism Research*, *3*(5), 226–236. https://doi.org/10.1002/aur.148

Westermann, G., & Mareschal, D. (2014). From perceptual to language-mediated categorization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), https://doi.org/10.1098/rstb.2012.0391

Wozniak, R., Leezenbaum, N. B., Northrup, J. B., West, K., & Iverson, J. (2016). The development of autism spectrum disorders: Variability and causal complexity: Development of autism spectrum disorders. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*, https://doi.org/10.1002/wcs.1426

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Tovar ÁE, Rodríguez-Granados A, Arias-Trejo N. Atypical shape bias and categorization in autism: Evidence from children and computational simulations. *Dev Sci*. 2019;e12885. https://doi.org/10.1111/desc.12885