SUPPLEMENTARY MATERIALS: Moya, Cristina. (2013). Evolved priors for ethnolinguistic categorization: A case study from the Quechua–Aymara boundary in the Peruvian Altiplano. *Evolution & Human Behavior*, 34(4) 2013, 265-272.

Raw data files are available at https://sites.google.com/site/cristinasolermoya/

1. Demographic descriptions of the participants by study.

Table S1. Participant demographics by Study number

	N	age				% female
		mean	SD	median	range	
Study 1	178	15	17	7	3-80	57
Study 2	73	16	15	9	3-62	55
Study 3	176	14	13	9.5	3-80	55
Study 4	138	25	21	18	3-75	64

2. Social category cues by Study number

Table S2. Social category cues.

Each character was assigned a set of ethnolinguistic cues and a set of occupational cues. Only the pairs of ethno-linguistic labels marked with superscripts 1 and 2 were used in Study 2. Additionally, Study 2 was prefaced with the following introduction that analogized the context to the local setting; “There is a foreign country far away where many people live. They work doing various things. They know how to subsist in different ways. In this place people live in two communities; Necia where they speak Nituano, and Sorega where they speak Sabones. In three hours you can walk from one to the other.” Study 3’s introduction was limited to the first sentence of Study 2’s preface.

Linguistic		**Occupational**	
Character A	Character B	Character A	Character B
a) Study 1: Real world			
“Is from [ ] and speaks [ ]”		“Works [ ] and [ ]”	
Q'ello Chico /Quechua	Ancomarca/Aymara	as a math/science teacher	as a nurse/technician
Toquepane/Aymara	San Salvador/Quechua	selling vegetables/groceries	as a cobbler fixing shoes/sandals
Qanqo/Quechua	Pongone/Aymara	doing agriculture/shepherding	as a mechanic fixing cars/trucks
Sustia/Aymara	Qhealli/Quechua	doing agriculture of potatoes/quinoa	shepherding cows/sheep
b) Studies 2 & 3: Fictional			
“Is from [ ] and speaks [ ]”			
Sorwega[1]/Sorwego	Nuecia[2]/Nueco	works as a lacaquero	works as a memolero
Nituania/Nituano[2]	Lepal/Lepales	knows how to melear better	knows how to nosear better
Lordania/Lordano	Jibya/Jibio	works in the country doing pekin	works in the city doing charis
Gomalia/Gomali	Sabon/Sabones[1]	is paid to clinte	is paid to chepre

3. List of Traits

Table S3. List of traits.
The 'Type' column reflects the a priori ascriptions not necessarily the opinions of participants. Traits with asterisks* were used in Study 4.

Type	Abbreviation	Traits
Cultural	Music*	Likes [mipu / padu] music
	Food*	Likes eating [bate/mino]
	Clothing*	Wears clothing with [sachu / reka] designs
	Distribution Norm	Thinks [everyone should receive equal / those that pay more should receive more] qami
Coalitional	Friendship	Is friends with [Yesica / Zaida]
	Political*	Voted for [Juan/José] for mayor
Occupational	Book	Likes reading books by [Sartre / Camus]
	Activity	Likes helping his friends [building houses / doing fieldwork]

4. Choice of occupation types and trait types.

For the first three studies we tried to choose traits for which we could easily come up with novel trait alternatives that participants would not have heard of (e.g. such as musical styles). However, the activity and distribution norms were less novel. The former was meant to ensure that participants were willing to make-occupation based ascriptions since we expected that activity preference would map onto occupational choices. The distribution norm trait reflected a characteristic about which there was much local debate and was of theoretical interest.

When we simplified the procedure for Study 4 we chose only 4 of these traits, and restricted ourselves to the ones that were completely novel, and had seemed easiest for participants to understand. When we run the analysis for Studies 1-3 using the restricted set of traits used in Study 4, the first three Studies still do not show a consistent effect of age or socialization. This suggests that the choice of subset of traits for Study 4 did not drive the developmental effect.

For Study 1 we chose a wide array of occupational comparisons, some of which reflected non-market subsistence lifestyles (e.g. agriculture versus shepherding), some of which reflected market-integrated professions (e.g. nurse versus teacher), and some which straddled this divide (e.g. agro-pastoralism versus mechanic). There was a slight suggestion that the questions using occupational labels that straddled this market-integration divide motivated fewer ethnic inferences *(OR=.89, SE=.11, p=.36)*.
In order to reflect the diverse occupational comparisons used in Study 1 we worded the occupational labels in Studies 2 and 3 to imply that certain jobs were associated with market integration. For example, "being paid to clinte" or "being paid to chepre" connoted equivalent levels of market integration since they are salary based, perhaps comparable to the "teacher versus nurse" comparison in Study 1. Working "in the country doing pekin" versus "in the city doing charis" was intended to mirror the market nature of "doing agro-pastoralism versus mechanical work" in Study 1 by virtue of the

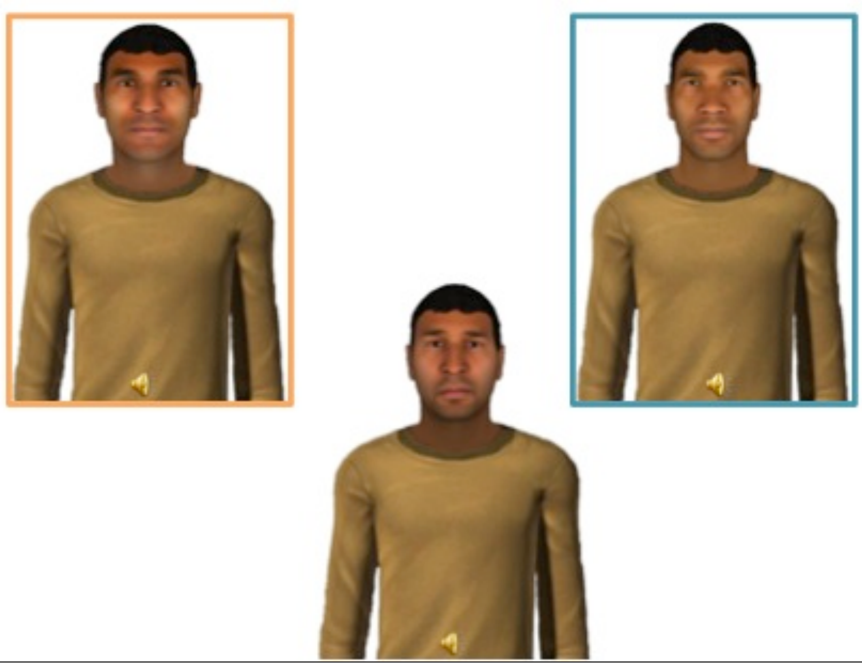
rural-urban distinction. We tried to do this in a various ways to minimize the risk of any particular framing driving the effect. We also tried to avoid lexicalized versions of categories (e.g. Quechua speaker) as these have been shown to increase inference rates (see Gelman & Heyman, 1999, *Psych. Sci.*, **10**, 489-493).

In Study 4 we stacked the deck as heavily as we could in favor of occupational predictions by choosing the Miner versus Agro-Pastoralist divide. This is a division that is common in town and structures a lot of economic variation, the former being wealthier, and more market integrated than the latter.

It is true that the forced choice nature of the design means that it is difficult to discern whether participants were motivated by a belief in the inductive potential of one category as opposed to the lack of inductive potential of the other. However, finding biases towards ethnic based predictions if the contrasting category was a less socially relevant one would have been less compelling evidence of the importance of ethno-linguistic units. Furthermore, it is difficult to come up with an evolutionary account of why individuals should have strong priors one way or another about occupational categories given that the degree of specialization that we see in market economies is relatively recently culturally evolved. Hirschfeld in his 1998 book *Race in the Making* does show evidence that children treat occupational categories as stable and possibly biologically inherited, though the fact that they are marked with uniforms makes the result difficult to interpret.

5. Stimuli for Study 4

Figure S1. Sample computer screen for Study 4. Audio icons would start each characters introduction when clicked.



6. Alternative analyses using various treatments of age

6.1 Age as categorical variable

Since we sampled heavily from children under 8 years of age we also show results with a binary split of age at 8 years (Table S4) and with 3 age category variables separating out adolescents from adults (Table S5). These should be more robust given that age and the socialization index (SI) deviated from normality. They also allow for non-monotonic effects of age. However, they also limit are of more limited statistical power and require choosing arbitrary cut offs when exploring the developmental trajectory. To control for the non-independence of participants' responses, the predicted probabilities in the following table are from logistic regression models including a random effects for participant. The results below are qualitatively similar to those using the SI as a predictor. The binary split analysis shows that individuals older than 8 years of age consistently make fewer ethnolinguistic inferences than those under 8, but only significantly so for Study 4. The ternary split shows no clear monotonic changes with age for Studies 1-3, and a clear decrement in ethno-linguistic inferences with increasing age category in Study 4, particularly driven by the oldest age category.

Table S4. Effect of age on probability of making an ethnolinguistic inference (binary split). Odds ratios and predicted probabilities from random effects logistic regression models predicting ethnolinguistic responses from categorical age variables (< 8 and ≥ 8). The younger category is the reference.

	age	OR	SE	p	n	Predicted probabilities (95% CI)	
Study 1	< 8	-	-	-	103	0.50	(0.46, 0.54)
	≥ 8	0.83	0.11	0.18	75	0.46	(0.41, 0.51)
Study 2	< 8	-	-	-	30	0.54	(0.44, 0.63)
	≥ 8	0.91	0.24	0.73	43	0.51	(0.42, 0.59)
Study 3	< 8	-	-	-	78	0.56	(0.49, 0.62)
	≥ 8	0.94	0.15	0.71	98	0.54	(0.49, 0.59)
Study 4	< 8	-	-	-	57	0.62	(0.54, 0.70)
	≥ 8	0.56	0.13	0.01	81	0.48	(0.41, 0.55)

Table S5. Effect of age on probability of making an ethnolinguistic inference (using three age categories). Odds ratios and predicted probabilities from random effects logistic regression models predicting ethnolinguistic responses from categorical age variables; young children (< 8), late childhood and adolesence (8≤ & ≤20), and adults (20 <). The youngest category is the reference.

	age	OR	SE	p	n	Predicted probabilities (95% CI)	
Study 1	< 8	-	-	-	103	0.50	(0.46, 0.55)
	8≤ & ≤20	0.89	0.15	0.48	37	0.47	(0.41, 0.54)
	20 <	0.78	0.13	0.14	38	0.44	(0.37, 0.51)
Study 2	< 8	-	-	-	30	0.54	(0.44, 0.63)
	8≤ & ≤20	1.23	0.36	0.47	24	0.59	(0.48, 0.69)
	20 <	0.62	0.20	0.13	19	0.42	(0.30, 0.54)
Study 3	< 8	-	-	-	78	0.56	(0.49, 0.62)
	8≤ & ≤20	0.88	0.16	0.47	64	0.52	(0.46, 0.59)
	20 <	1.07	0.23	0.75	34	0.57	(0.49, 0.66)
Study 4	< 8	-	-	-	57	0.62	(0.54, 0.70)
	8≤ & ≤20	0.86	0.32	0.68	15	0.58	(0.43, 0.74)
	20 <	0.51	0.12	0.004	66	0.45	(0.37, 0.53)

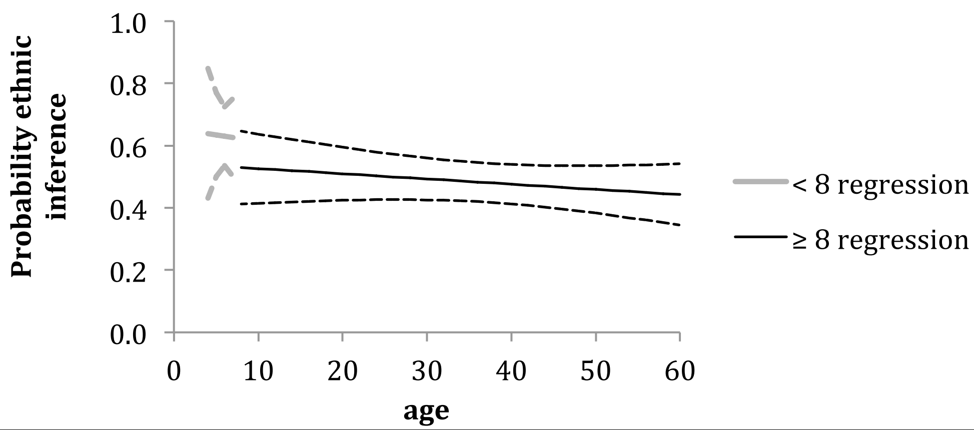
6.2 Separate regressions by age category

Alternately we can run the regressions separately for different age categories to test whether the effect of age is different in children versus adults. Table S6 does suggest that the decrement in ethnic inference rates with age is slightly larger in the under 8 years old category than in the older category. However, none of the linear effects of age within any age category or within any study are significant. Figure S2 illustrates why this is the case for Study 4, despite our robust age effects for that study reported above – there is a large jump in the predicted rates of ethnic predictions from age 7 to age 8 that is not captured in either regression using this method.

Table S6. Effect of age on probability of making an ethnolinguistic inference (separate regressions by age category). Odds ratios and their standard errors are from random effects logistic regression models predicting ethnolinguistic responses from a continuous age variable. Separate models were fit for each study for participants in the 2 age categories (< 8 and ≥ 8).

	age	OR	SE	p	n
Study 1	< 8	0.96	0.08	0.64	103
	≥ 8	1.00	0.01	0.39	75
Study 2	< 8	0.96	0.136	0.78	30
	≥ 8	0.99	0.01	0.38	43
Study 3	< 8	0.94	0.0086	0.51	78
	≥ 8	1.01	0.008	0.53	98
Study 4	< 8	0.98	0.19	.92	57
	≥ 8	0.99	0.01	0.33	81

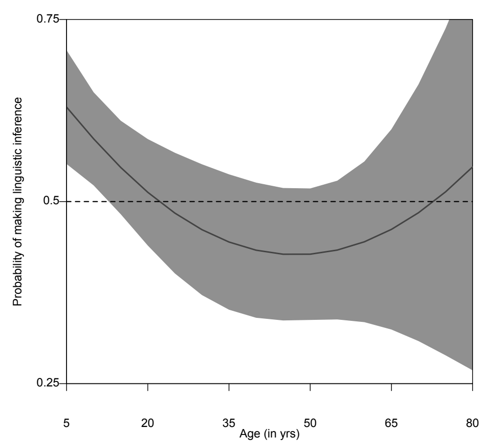
Figure S2. Study 4. Predicted probabilities of ethnic inference as a function of age for under and over 8 year olds. From random effects logistic regressions run separately for over and under 8 year olds. Dotted lines denote 95% CIs.



6.3 Linear and squared age terms

We also ran models using raw age as a predictor, and testing for non-linear effects by adding an age squared term. We report more broadly about these in the next section, but here we limit ourselves to Study 4 where in fact we see a marked developmental change. We find that the models allowing for an age squared term fit slightly better (have lower AIC scores) than those with linear age terms alone. The linear age term is significantly negative in all models in either case. However, allowing the squared term allows us to investigate the full developmental trajectory (see Figure S3). Confidence intervals from the model are very wide for older individuals, of which we have very few, but the predicted increase in ethnic inferences beyond 50 years of age might reflect a secular trend wherein ethno-linguistic categories are becoming less important organizers of social life.

Figure S3. Predicted probability of making ethnic inference in Study 4. From model with age and age squared terms. Grey areas represent 95% CIs.



6.4 Deriving the best fit Socialization Index (SI)

For the purposes of this paper we are not interested in this potential upward trajectory in later adulthood that is likely to reflect historical changes rather than cognitive development. Therefore, we develop a Socialization Index (SI) which is simply a negative exponential function of raw age. The negative exponential function is asymptotic and so collapses all variation at higher values of age.

However, we have little *a priori* reason to expect a particular negative exponential function since we know relatively little about this developmental trajectory. For this reason we compare models with varying negative exponential functions and choose the best fit one as determined by the lowest AIC score. We test the SI's derived from the following function:

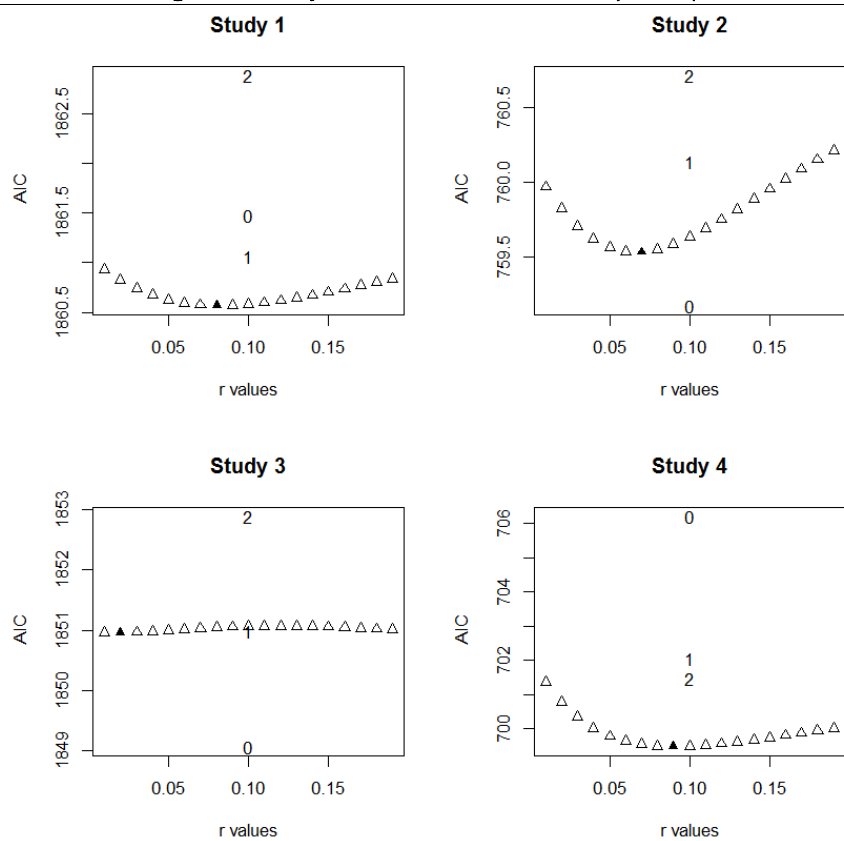
$$SI = 1 - e^{-r*age}$$

for values of r ranging from 0.01 to 0.2 in 0.01 unit increments.

For reference, most AIC scores of the SI's tested were lower those for the baseline model without age as a predictor, a model with a linear raw age term, and of a model with age and age squared as predictors. AIC's for these alternate models are shown as 0, 1, and 2, respectively, in Figure 2 alongside AICs for various SI's as a function of values of r.

Running these models we find that negative exponential models fit best for Study 1 and Study 4, while baseline models fit best for Studies 2 and 3. For Study 1 the best fit negative exponential is for r=0.08 and for Study 4 it is when r=0.09 (see dark triangles in Figure S4). Given that the SI function has an asymptote of 1, we can set SI to 0.5 to calculate at what age we would expect participants to reach half-way between initial and adult rates of making ethnic inferences. For Study 1 this expected "half-way" point would be 8.7 years old while, while for Study 4 it would be 7.7.

Figure S4. AIC scores for various models of age. Triangles indicate AIC scores for models using Socialization Indices with varying r-values in the negative exponential function. The filled in triangle denotes the r value with the lowest AIC score among these. The numeric markers denote AIC scores for a baseline model with different orders of age terms; without predictors (0), a model with a linear age term as a predictor (1) and a model with a linear and squared age terms (2). The position at r=0.1 for these is not meaningful and is just meant to allow easy comparison.



7. On the non-independence of observations

The primary goal in using random effects models was to control for the non-independence of observations across questions within participant. However, we also used this method to conservatively account for the possibility that the same individual participated multiple times in any one of our studies. This was suspected to be a problem among children especially, who enjoyed participating enough that they tried to do so multiple times. This was a larger problem in the earlier field season when Studies 1-3 were run, than the later field season when Study 4 was run and the experimenters were better acquainted with the potential participants. In going back through the dataset I collapsed observations of individuals who had similar names and ages and gave them a single id number. This was used as the individual cluster level in the random effects models and is a conservative measure, as it decreases the independence of observations, thus increasing the standard error estimates.

8. On the absence of sex differences

We did not expect, nor did we find, consistent sex differences in how much participants relied on linguistic versus occupational cues to make inductive inferences. Controlling for the socialization index, being female slightly, but not significantly, increased the likelihood of making ethno-linguistic predictions across the first three studies (OR=1.14, SE = 0.12, p=0.19) and decreased it non-significantly in Study 4 (OR=0.85, SE = 0.20, p=0.49). As such we did not include sex as a predictor in the following models.

We also found no difference in ethnolinguistic inference rates as a function of the gender of the illustrated characters. Across Studies 1-3 the effect of the characters being female on the probability that participants made ethnolinguistic inference on their round was OR=1.07, SE=.08, p=.314. As such, we do not expect that restricting stimuli to male characters only in Study 4 had any meaningful on the results. If anything, using male characters should have reduced overall rates of ethnolinguistic given results from Studies 1-3, but in fact we saw higher rates. Finally, we see no interaction effects of participant and triad character gender on rates of ethnolinguistic inference (either for children or adults).

9. The effect of trait type and occupation type on rates of ethnolinguistic inference.

We tested models with occupation type – dummy coded as representing a market-integration divide (1) or not (0) – and type of novel trait – dummy coded as expected to be cultural (1) or not (0) as predictors. We expected occupations that differed in terms of market integration would motivate people to make more occupation based inferences, thus reducing further their reliance on language-based inference. We predicted questions about cultural traits – as outlined in Table S3 – to have the opposite effect by motivating a greater reliance on ethno-linguistic inference. However, these models did not fit the data better than simpler model with no predictors or with a main effect of Socialization Index (SI), see Table S7. Table S7 compares 8 different models for each of the first three studies. These include baseline models with no predictors and with SI as a predictor, and models with additional parameters of either trait type or occupation type. Models with interactions of trait and occupation type with SI are also tested to determine if the differentiation of occupations and traits develops as the child ages. Models numbered 6 and 7 in Table S7 suggest that participants were not particularly sensitive to whether the occupation comparison stradded the market integration divide. While the effects are in the expected direction of occupations signaling differences in market integration reducing the probability of making an ethnic inferences (negative logit coefficients), these effects are are weak at best. The low AIC weights of all models with occupation and trait type and the small effect sizes for each of these predictors indicate that simpler models explain the data better.

A possible exception is seen in Study 3 where Model 5 fits relatively well and the interaction term suggests that ethnic inferences about cultural traits decrease with socialization. However, overall a consistent pattern does not emerge. The effect of trait type on rates of language-based inferences did not map neatly onto our expectation that cultural traits would foster more linguistic predictions. In large part this was because people inferred that language-category membership would also be predictive of friendship and mayoral choices which were posited to be “coalitional” (see Figure S5). This may be because regional assortment in communities that are relatively linguistically homogenous lead to linguistically homogenous social networks. In accordance with our hypothesis about occupational traits, activity and book preferences motivated the fewest ethnic inferences.

Table S7. Random effects models fit predicting language-based inference from socialization index (SI), trait type (dummy-coded as 1 for cultural traits), and occupation type (dummy-coded as 1 for occupations associated with the market integration divide). Model comparisons are done separately for a) Study 1, b) Study 2, and c) Study 3. Note: In these models SI = $20*(1 - e^{(-0.2* age)})$, but other treatments of age did not qualitatively change the results.

a) **Study 1**

Model	**Logit coef.**	**SE**	**p**	**AIC**	**AIC weight**
1				1861.47	0.17
constant	-0.07	0.07	0.30		
2				1860.88	0.22
SI	-0.03	0.02	0.11		
constant	0.44	0.32	0.18		
3				1862.78	0.09
trait type	0.09	0.11	0.41		
constant	-0.12	0.09	0.18		
4				1862.17	0.12
SI	-0.03	0.02	0.11		
trait type	0.09	0.11	0.40		
constant	0.39	0.33	0.23		

5				1864.04	0.05
SI	-0.03	0.03	0.31		
trait type	0.28	0.54	0.60		
trait X SI	-0.01	0.03	0.72		
constant	0.30	0.42	0.47		
6				1862.63	0.09
occupation type	-0.12	0.13	0.36		
constant	-0.04	0.07	0.59		
7				1862.03	0.13
SI	-0.03	0.02	0.11		
occupation type	-0.12	0.13	0.36		
constant	0.47	0.33	0.15		
8				1861.81	0.14
SI	-0.02	0.02	0.42		
occupation type	0.79	0.62	0.21		
occu. X SI	-0.06	0.04	0.14		
constant	0.24	0.36	0.50		

b) Study 2

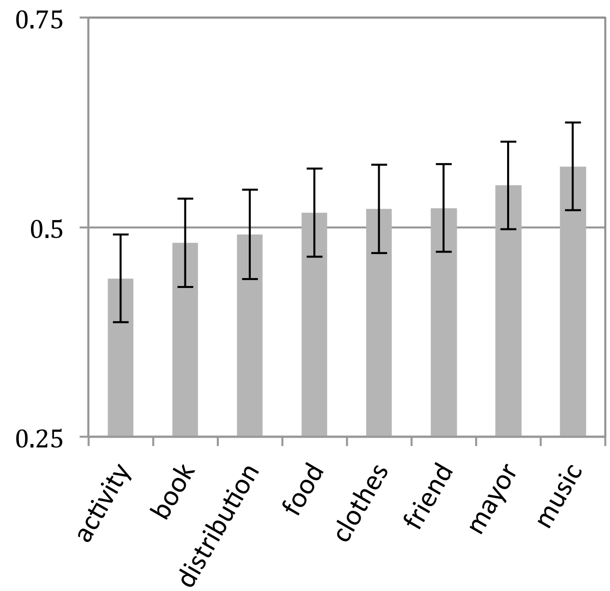
Model	**Logit coef.**	**SE**	**p**	**AIC**	**AIC weight**
1				1849.07	0.35
constant	0.19	0.08	0.02		
2				1850.99	0.14
SI	-0.01	0.03	0.78		
constant	0.31	0.45	0.49		
3				1850.34	0.19
trait type	0.10	0.11	0.39		
constant	0.14	0.10	0.17		
4				1852.26	0.07
SI	-0.01	0.03	0.78		
trait type	0.10	0.11	0.39		
constant	0.26	0.45	0.57		
5				1853.38	0.04
SI	-0.03	0.03	0.44		
trait type	-0.49	0.64	0.44		
trait X SI	0.04	0.04	0.35		
constant	0.55	0.55	0.31		
6				1851.06	0.13
occupation type	-0.02	0.13	0.91		
constant	0.19	0.09	0.03		
7				1852.98	0.05
SI	-0.01	0.03	0.78		
occupation type	-0.02	0.13	0.91		
constant	0.31	0.45	0.48		
8				1854.23	0.03
SI	-0.02	0.03	0.55		
occupation type	-0.64	0.73	0.38		
occu. X SI	0.04	0.04	0.39		
constant	0.47	0.48	0.33		

b) Study 3

Model	Logit coef.	SE	p	AIC	AIC weight
1				757.39	0.22
constant	0.08	0.13	0.53		
2				758.40	0.13
SI	-0.04	0.04	0.32		
constant	0.78	0.71	0.28		
3				758.25	0.14
trait type	0.19	0.18	0.29		
constant	-0.01	0.16	0.93		
4				759.25	0.09
SI	-0.04	0.04	0.32		
trait type	0.19	0.18	0.28		
constant	0.69	0.72	0.34		
5				757.63	0.20
SI	0.01	0.05	0.80		
trait type	2.07	1.01	0.04		
trait X SI	-0.11	0.06	0.06		
constant	-0.23	0.87	0.79		
6				758.96	0.10
occupation type	-0.14	0.21	0.52		
constant	0.12	0.14	0.41		
7				759.98	0.06
SI	-0.04	0.04	0.32		
occupation type	-0.14	0.21	0.52		
constant	0.82	0.72	0.26		
8				759.98	0.06
SI	-0.07	0.05	0.15		
occupation type	-1.73	1.15	0.13		
occu. X SI	0.10	0.07	0.16		
constant	1.22	0.78	0.12		

SUPPLEMENTARY MATERIALS: Moya, Cristina. (2013). Evolved priors for ethnolinguistic categorization: A case study from the Quechua–Aymara boundary in the Peruvian Altiplano. *Evolution & Human Behavior*, 34(4) 2013, 265-272.

Figure S5. Effect of trait type on linguistic inductive inference rate. Predicted probabilities of making a language based assessment from random effects models. Bars reflect 95% CIs.



10. On the absence of order effects.
There may have been unintentional order effects (either experimenter induced or as participants familiarized themselves with the protocol) that pushed participants answers towards more or fewer ethnic inferences as they proceeded with the task. In order to test for these we examine the effect of question order within subjects on their responses. Table S8 shows no linear effect of question order on the probability of making ethnic predictions across participants of all ages in any of the studies. Table S9 shows that order effects are also absent for children under 8 years of age. To examine the children's pattern in more detail Table S10 shows the proportion of ethnic responses for each question position. No discernible pattern is found across studies and X^2 tests reveal no significant deviations from the expected distribution of responses for each study.

Table S8. Effect of increasing question order on probability of making an ethnic inference – All participants. These results are from random effects logistic regression models that control for participant id. In Studies 1-3 question order took integer values from 1-8, while in Study 4 they took values from 1-4.

	OR	SE	p
Study 1	0.99	0.02	0.66
Study 2	1.04	0.04	0.28
Study 3	0.98	0.02	0.42
Study 4	1.05	0.09	0.54

Table S9. Effect of increasing question order on probability of making an ethnic inference – Participants under 8 y.o. only. These results are from random effects logistic regression models that control for participant id. In Studies 1-3 question order took integer values from 1-8, while in Study 4 they took values from 1-4.

	OR	SE	p
Study 1	0.98	0.03	0.47
Study 2	1.09	0.06	0.17
Study 3	1.01	0.04	0.65
Study 4	1.05	0.14	0.70

Table S10. Proportion of responses that are ethnic inferences by question order. - Participants under 8 y.o. only. Note missing values indicate that the participant refused or did not answer a particular question. Study 4 only included 4 questions.

	question order									
	1	2	3	4	5	6	7	8	Total	X^2 (p)
St 1										
% eth	53.9	47.0	48.0	52.5	51.1	50.6	53.9	41.6	49.9	4.5 (.72)
N	102	102	102	99	92	91	91	89	768	
St 2										
% eth	53.3	46.7	40.0	58.6	44.8	65.5	58.6	58.6	53.2	6.2 (.51)
N	30	30	30	29	29	29	29	29	235	
St 3										
% eth	52.6	56.4	60.3	60.3	38.7	60.0	52.8	63.4	55.5	13.0 (.07)
N	78	78	78	78	75	75	72	71	605	
St 4					-	-	-	-		
% eth	50.9	73.2	60.0	58.9	-	-	-	-	60.81	6.0 (.11)
N	55	56	55	56	-	-	-	-	222	

11. Mean ethnic inference score as dependent variable

Instead of running random effects logistic regressions, we can give each individual an ethnic inference score which is the average of their answers on the questions for which they made an ethnic- (1) or an occupation-(0) based inference. We do not count non-answers in either direction. Scores can be between 0 and 1 with higher values indicating a greater proportion of an individual's responses were based on the ethno-linguistic cue. Table S11 and Figure 2 show that the results from this analysis are comparable to those of the RE logistic regressions.

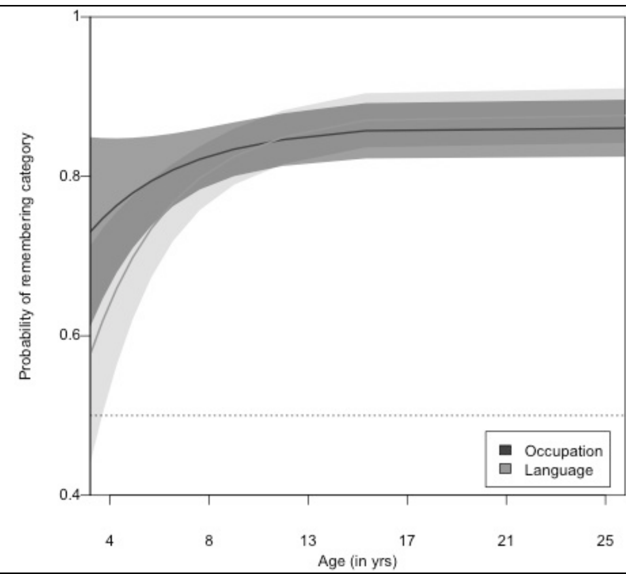
Table S11. Mean proportion of ethnic inferences per individual by age category and study. Standard errors in parentheses.

		<8	8 - 20	20<
Study 1	n	102	37	38
	mean (SE)	0.5 (0.02)	0.47 (0.03)	0.42 (0.03)
Study 2	n	30	24	19
	mean (SE)	0.53 (0.04)	0.58 (0.05)	0.45 (0.07)
Study 3	n	78	64	34
	mean (SE)	0.55 (0.03)	0.52 (0.03)	0.57 (0.03)
Study 4	n	56	15	65
	mean (SE)	0.61 (0.04)	0.58 (0.05)	0.44 (0.04)

12. Effect of Socialization Index on memory for occupational and ethnic information in Study 4

In Study 4 linguistic information was provided through more channels than was the occupational information (i.e., through the character's speech and his labeling himself). Therefore, it might be that children's preferential use of language for inductive inference was driven by a memory bias due to this redundancy. If this were the case then the socialization effect should have gone away once we controlled for memory checks. However, Figure S6 shows the predicted probabilities of remembering each of the characters' occupational and language category membership on the first try from RE logistic regression models using SI as a predictor. It suggests that if anything, children remembered the occupational information better than the linguistic information.

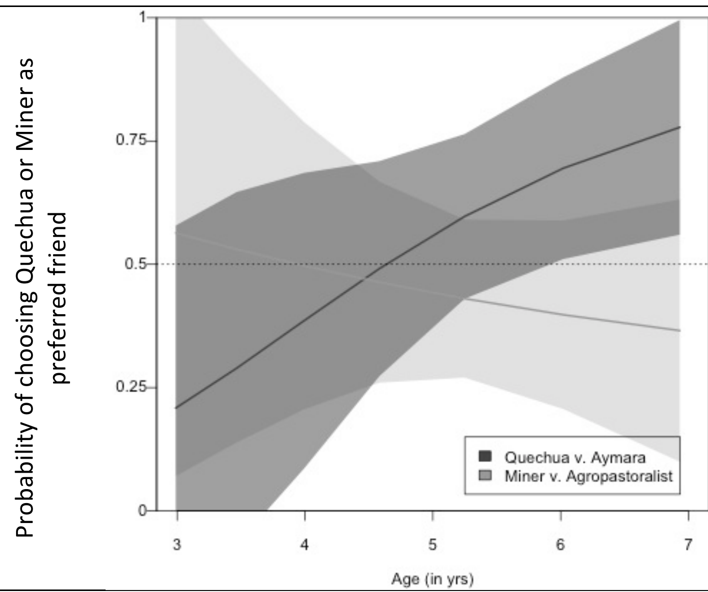
Figure S6. Probability of remembering category labels in Study 4.
Best fit random effects logistic regression models predicting probability of remembering label as a function of socialization index. The 95% confidence intervals for the models are estimated using the Delta-method of standard error estimation. Participants ranged from children three years old to adults in their 70's. Models were fit using the full range of the socialization index, but restricted age ranges are plotted below for ease of interpretation and to improve resolution.



13. Effect of Socialization Index on preference for specific category members

In a separate study (Study 5) we used the same acoustic and visual stimuli as in Study 4 to explore explicit biases in hypothetical friendship choices. We find no strong explicit bias for either linguistic or occupational category among 35 children aged 3-7. However, a preference for ingroup language speakers (Quechua speakers) starts to emerge in the older children in this sample around age 6 (see Figure S7). If we run the same analysis for Study 4 restricting our sample to equivalently aged children less than 8 years of age we see the same trajectory that we report in other analysis in the paper *(OR = .47, SE=1.7)* – that is while ethno-linguistic based inferences are decreasing from age 3-7, explicit linguistic in-group biases are emerging.

Figure S7. Probability of preferring Quechua-speaker or miner using speech cue stimuli. (Study 5)
Best fit random effects logistic regression models predicting probability of preferring the Quechua-speaker over the Aymara one (dark grey), and the miner over the agropastoralist (light grey), as a function of socialization index. The 95% confidence intervals for the models are estimated using the Delta-method of standard error estimation. Participants were children from 3 to 7 years of age. Models were fit using the the socialization index, but ages are plotted below for ease of interpretation.

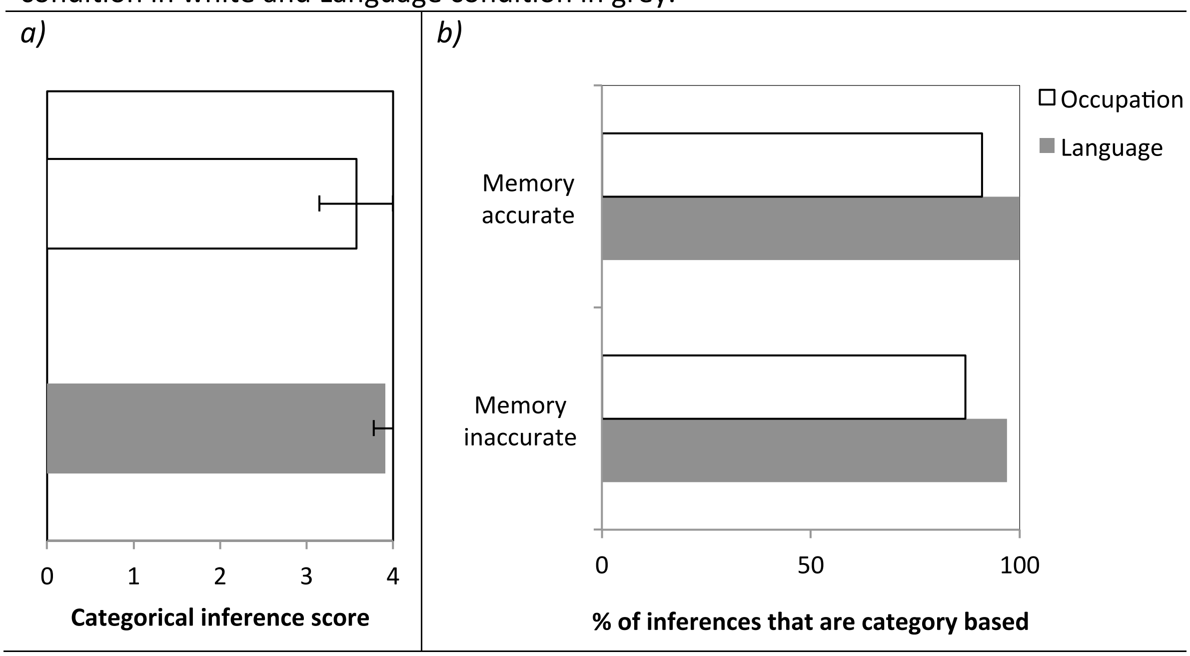


14. Inductive inference rates when task is not forced choice (Study 6)

In a separate study (6) we recruited 66 children between the ages of 4 and 8. We used a similar triad task as in Studies 1-3 but only asked participants to make predictions about 4 real world and functionally relevant traits: being good at making medicinal teas or not, being good at taking care of kids or not, being good at helping with homework or not, and being good at soccer or not. More importantly, we changed the task to a between-subjects design wherein some participants were randomly assigned to the occupation condition and the other half were assigned to the language condition. The former were told that the characters were either miners or agro-pastoralists, while the latter were told that the characters were Quechua or Aymara speakers. Two out of the three characters therefore shared a category membership, while the other one did not. This task also included a memory check after each triad was introduced.

Figure S8 shows that the vast majority of responses were inferences based on the characters' sharing the same category membership. This is true for both conditions regardless of the participants' performance on the memory check. Admittedly, this is not surprising since the non-category member did not share any particular characteristic with the target character. Inferences based on category membership were slightly more common in the language condition *(t=1.52(64), p=.13)*. This is consistent with our results from Studies 1 and 4 with real world categories. However, the fact that only 15 out of 262 responses were not predictions based on joint category membership suggests a ceiling effect, making strong inferences about condition effects difficult. On the other hand, this result does suggest that children were capable of engaging a triad task very similar to the others in this paper, and of giving decidedly non-random responses. It is likely that when 2 conflicting categories were used in a single triad task the inductive potential of each category brought responses closer to 50% for each category.

Figure S8. Reliance on category based inferences in Study 6. Participants' category based inference scores (0-4) with 95% CI are shown in *a*. The percent of inferences that were category-based by performance on memory check are shown in *b*. Occupational condition in white and Language condition in grey.



15. Randomization and counter balancing procedures

Several parameters were randomized and counterbalanced. In Studies 1-3 the gender of the triad illustrations alternated from all female to all male between questions. To avoid primacy effects or side biases, the order in which the characters' linguistic and occupational information was presented, the side of the linguistically similar test character, and the pairing of a specific occupational category with a specific language category were counterbalanced across scenarios within subjects. The orders in which participants were asked about traits, and in which the specific social category labels were used were randomized independently for each participant. For Study 4 randomizations and counter-balancing analogous to those for the other studies were carried out when relevant by switching between slide presentations on the computer.