Disentangling Language and Education Effects on False Belief Understanding: Evidence From Homesigners, Signers, and Unshooleo Spanish Speakers

Deanna L. Gagne (deanna.gagne@uconn.edu)
Department of Psychology, 406 Babbridge Road, Unit 1020
Storrs, CT 06269-1020 USA

Marie Coppola (marie.coppola@uconn.edu)
Departments of Psychology and Linguistics, 406 Babbridge Road, Unit 1020
Storrs, CT 06269-1020 USA

Abstract
To investigate the contributions of language and education to Theory of Mind understanding, three Nicaraguan groups were tested using a minimally verbal protocol in which they themselves experienced a false belief instead of being told of one. We also assessed the relationship of executive function abilities to false belief performance. Homesigners, who have no linguistic community, did not succeed on either the False Belief or executive function tasks. Nicaraguan Sign Language users, who have educational experience and are part of an emerging linguistic community, performed the best on executive function, though less well on false belief, than Spanish speakers who have little to no education. This study showed that: without a language community, succeeding on either task is difficult; executive function may not be as tied to false belief performance as previously believed; and education may play a greater role in executive function success than language does.

Keywords: Theory of Mind; Simulation Theory; false belief; social cognition; executive function; homesign; Nicaraguan Sign Language; language community; education.

Introduction
Theory of Mind (ToM), the ability to understand that others have beliefs and desires different from one’s own, and that those beliefs and desires may influence their behavior (Wimmer & Pernier, 1983, among others), is a foundational aspect of human social cognition. Whether this ability is primarily driven by language development, life experience, or social interactions is actively debated.

Proponents of so-called Simulation Theories argue that life experience and social interactions are sufficient to scaffold mature ToM abilities, without the need for language (e.g., Gordon, 1986; Goldman, 1992). On this view, one’s own experiences provide the foundation for ToM by serving as a template for understanding how others will behave in a given similar situation; these templates are updated as one gains more life experience. Language can play a part, because it provides a way to interact more, and to thus update one’s templates more efficiently, but it is the reflection on one’s own experiences that primarily supports the understanding of other minds.

Other researchers argue that language development, and even specific linguistic structures, are essential to the child’s development of a mature ToM. Mental verbs (e.g., Howard et al., 2008; Gola, 2012) and complement structures (e.g., deVilliers & Pyers, 2002) have been associated with theory of mind abilities. This is not simply based on linguistic complexity: Hale & Tager-Flusberg (2003) found that training on sentential complements, but not relative clauses, improved children’s ToM performance relative to pre-training levels. Those advocating language-based theories argue that some aspect of language is central to normal ToM development.

We cannot easily disentangle these factors in typically developing children because language and social abilities develop simultaneously. By 4-5 years of age, children have amassed enough social experience and language to successfully navigate the gold standard task for assessing ToM: false belief (FB) (e.g., Wimmer & Pernier, 1983).

Working with deaf individuals who have not acquired conventional language, and instead use gestural systems, called homesign, to communicate (Goldin-Meadow, 2003) enables us to distinguish these factors in ways that other groups cannot. Despite their lack of conventional linguistic input, homesigners who continue to use their gesture systems as a primary language as adults are not socially impaired; they enjoy relatively typical social interactions with their hearing families, friends, and neighbors. While not full languages, mature homesign systems exhibit a range of linguistic properties (e.g., the grammatical relation of subject, Coppola & Newport, 2005).

One of the main advantages of research with recently emerging languages is the opportunity to measure the effects of absent or atypical linguistic input, and the linguistic and cognitive benefits conferred by participating in a linguistic community. Neither of the two groups of deaf participants studied here, Homesigners and Cohort 1 Signers of Nicaraguan Sign Language (NSL) (the first group of signers to begin creating the language in the 1970s), had access to linguistic input transmitted vertically, that is, from a pre-existing language model. However, Cohort 1 Signers did engage in language genesis with their peers (horizontal transmission) (Senghas et al., 2005). NSL Signers of all cohorts interact with many other users who use the system as a primary language, i.e., members of the deaf community in Managua. Though homesigners rely on their gesture systems for their entire lives, the hearing people around them use it only with the homesigner.
Further, each homesigner’s mother does not share the system with the homesigner, despite having used it over a long period of time (Carrigan & Coppola, 2012). By these criteria, the homesigner, hearing family members and friends who use homesign, and their patterns of interaction do not constitute a linguistic community (even a small one). Empirical and computational evidence suggests that the rich interconnections among users that characterize typical sociolinguistic communities are essential for developing some aspects of linguistic structure, e.g., a conventionalized lexicon (Richie et al., 2014). Thus, we compare Homesigners and NSL Cohort 1 signers to see the effects of having a linguistic community on ToM abilities.

Signers who were exposed to sign language via the Deaf community before 10 years of age performed better on ToM tasks than those who entered after age 10 (Morgan & Kegl, 2006). Cohort 2 signers, who represent a more developed stage of NSL, outperformed Cohort 1 signers (who were their linguistic models) on a change-location FB task (Pyers & Senghas, 2009). Success in both cohorts was related to the number of mental verbs used to describe short videos. On a task similar to the current task, Cohort 2 outperformed Cohort 1 overall (Pyers, 2005), with Cohort 1 signers having moderate success on Appearance/Reality items. Could this success be due to participation in a linguistic community, or simply a consequence of fundamental human abilities to observe and reflect?

Like most deaf children born to hearing parents who do not know a sign language, all NSL signers in past and present studies gained access to their linguistic community via educational settings. Thus, we cannot separate having a linguistic community from education in either NSL signers (who have both) or Homesigners, who have neither. We introduce Unschooled Spanish Speakers, who have full access to a language community, but have little to no education, to help us disentangle these factors.

We studied three extremely rare and understudied groups in Nicaragua to distinguish the contributions of language (i.e., linguistic community) and educational experience to ToM development (FB understanding) and executive function (inhibitory control). In Study 1 we induced each participant to unwittingly hold a false belief and then asked them to predict the behavior of another person who encountered the very same false belief. We predicted that Homesigners, who do not participate in a linguistic community, and thus are unlikely to have developed the linguistic structures that support FB understanding, would not succeed. Prior results suggested that some NSL Signers would succeed. Given that the Spanish speakers are adults with typical developmental histories, we expected them to succeed.

### Study 1: Theory of Mind

#### Participants

All participants were Nicaraguan: 4 Homesigners (HS), 6 signers from Cohort 1, representing NSL’s earliest users (NSL Signers), and 5 Unschooled Spanish Speakers (USS) who, like homesigners, had little to no education, but who, like the NSL signers, were part of a linguistic community (Table 1). Four of the five unschooled hearing participants were full-time agricultural workers; the fifth worked making tortillas for a family business. Their lack of education primarily resulted from economic restrictions and the distance to the nearest school.

#### Materials and Procedure

In order to study the effects of language on false belief and executive function abilities, we sought an alternative to traditional FB tasks, in which experimenters provide task instructions, content, and prompts using language, and in which participants respond verbally. We also avoided tasks that required experience with activities typical of middle-class home and educational contexts. For example, the picture-completion task used with NSL signers in Pyers & Senghas (2009) was not successful with Nicaraguan homesigners, and may also present issues for unschooled hearing Nicaraguans, who similarly may not appreciate the convention of sequentially ordered pictures representing events in a narrative (as in storybooks). In our design, we strove to eliminate language from both the task instructions and task demands, and therefore used an experiential FB task adapted from Pyers (2005), described below.

We gave each participant firsthand experience with Appearance-Reality (A/R) and Unexpected Contents (UC) false belief situations. They then participated in a prediction game in which they earned money for making correct predictions. We describe the procedure in great detail because the incremental, implicit understanding of the task instructions, and how participants should respond, are essential to our commitment to a minimally verbal procedure that fairly assesses the theory of mind abilities of homesigners in particular. Table 2 summarizes the three phases and 14 trials that each participant saw, first for the experience condition, then again during the prediction condition.

---

1 No secondary school was then available to NSL signers; they attended elementary school until ~16 years of age (completing 6th grade), and some then attended vocational school. Furthermore, a typical school day in Nicaragua tends to be of a ½ day format, due to a lack of resources in the school system. Thus, the reported mean is not comparable to completing some high school.

<table>
<thead>
<tr>
<th>Group</th>
<th>Age (mean, range)</th>
<th>Linguistic community</th>
<th>Educational experience (mean, range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homesigners</td>
<td>31.5y (26–35y)</td>
<td>No</td>
<td>Very little (0.5y, 0–1.5y)</td>
</tr>
<tr>
<td>NSL Cohort 1</td>
<td>41y (35–45y)</td>
<td>Yes</td>
<td>Yes¹ (10.5y, 6–13y)</td>
</tr>
<tr>
<td>Unschooled Spanish Speakers</td>
<td>28.5y (19–39y)</td>
<td>Yes</td>
<td>Very little (0.7y, 0–3y)</td>
</tr>
</tbody>
</table>

---

*¹ No secondary school was then available to NSL signers; they attended elementary school until ~16 years of age (completing 6th grade), and some then attended vocational school. Furthermore, a typical school day in Nicaragua tends to be of a ½ day format, due to a lack of resources in the school system. Thus, the reported mean is not comparable to completing some high school.
Phase 1: Stickers: The goal of Phase 1 was twofold. The experience condition familiarized participants with the process of choosing items from an array, and (nonverbally) demonstrated that a choice on a particular trial may be obvious (such as in an array of three originally identical stickers, one pristine, and the other two crumpled or ripped), or a choice might be based on preference (e.g., two different-colored smiley face stickers). In the prediction condition, the sticker phase provided the understanding that: 1) sometimes it is easy to predict someone else’s behavior (obvious choice trials), 2) sometimes it is hard (individual preference trials), and 3) correct predictions earn them a small monetary reward (5 Córdobas per correct prediction (max. 70 Córdobas or US$2.75) across all trials (a healthy incentive given typical local incomes).

Table 2. The phases and individual trials that each participant saw, once as an experiencer and then once as a predictor of a confederate’s choices. All relevant ordering possibilities were counterbalanced.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Trial</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stickers</strong> (familiarization with procedure)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Stickers – Obvious choice</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Stickers – Obvious choice</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Stickers – Individual preference</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Stickers – Individual preference</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Stickers – Obvious choice</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Stickers – Individual preference</td>
<td></td>
</tr>
<tr>
<td><strong>Appearance/Reality</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Cookies – 1, 2, 4 cookies on plates (Plate with 4 cookies is fake) (Experience False Belief)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Cookies – 1, 2, 4 cookies on plates (Plate with 4 cookies is fake) (Test Knowledge)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Tool matching – Pen with paper</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Tool matching – Pitcher with cup</td>
<td></td>
</tr>
<tr>
<td><strong>Unexpected Contents</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Tool matching – Matchbox (containing key) (Experience False Belief)</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Tool matching – Pen with notebook</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Tool matching – Pitcher with mug</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Tool matching – Matchbox (containing key) (Test Knowledge)</td>
<td></td>
</tr>
</tbody>
</table>

Phase 2: Appearance/Reality: In the A/R phase, the participant saw three plates holding 1, 2, and 4 cookies. Unbeknownst to the participant, the four “cookies” were very convincing ceramic composite replicas. The experimenter encouraged each participant to indicate the “best” plate. For the homesigners, this was done by pointing at the participant, and then to the three plates of cookies, followed by a thumbs-up gesture with a questioning look. All participants chose the plate with four cookies during their experience phase; when encouraged to try a cookie, they discovered that the cookies were not real. The cookies were returned to their original array and the question or gestures were repeated, this time to test the participant’s knowledge that the cookies were fake (thus the plate with 2 cookies should be chosen as the “best”).

Phase 3: Unexpected Contents: In the UC phase, the participant was presented with one of two arrays: paper, a glass, a lock, and a candle or a notebook, a mug, a lockbox and a box of cigarettes. The participant was then presented with a series of tools and was asked to indicate which object in the array each tool is used with. First, the participant was presented with a pen, and had to match it with an object (paper or notebook). Upon choosing the paper, the participant was asked to make a mark on the paper. The participant was then presented with a pitcher of water and asked to match it to an object (the glass or mug). Upon choosing, the participant was asked to pour water into the vessel. Third, the participant was presented with the matchbox (containing a key, but no matches), and again asked to match it to an object (in the False Belief trial, the participant should choose the candle or cigarettes, but in the knowledge trial, the participant should know that the matchbox contains a key and choose the lock or lockbox). Upon choosing the candle or cigarettes, the participant was encouraged to light the candle or a cigarette, and subsequently discovered that the matchbox contained a key, not matches. The array was then switched (i.e., paper to notebook, etc.) but the three tools (pen, pitcher, and matchbox) were left in view of the participant, so he or she could see that no one, including the experimenter, touched them. The entirety of Phase 3 was repeated, matching tools to their objects; the key trial was the matchbox trial, during which the participant should demonstrate his or her knowledge that the matchbox contained a key (instead of matches), and thus should be matched with the lock/lockbox, not the candle/cigarettes. In the prediction phase, the first two tools (pen, pitcher of water) set the participant up for the tool-matching paradigm, while the third (matchbox containing a key) presented a false belief, then a test of knowledge.

To reiterate, after experiencing the entire task, and more importantly, directly experiencing the false beliefs, each participant participated in all trials again, but this time as a predictor of another’s choices— a confederate who was a member of the research team but who had not previously participated in any aspect of this task. The participant indicated the item the confederate would choose, before the confederate actually made a selection, by marking a set of laminated sheets with images of all the arrays.

Results

A participant had to correctly predict the confederate’s choices in both the A/R and UC phases to pass the task overall (see Figure 1, white and black bars, for proportion of participants passing each subtest). None of the Homesigners, who lack a linguistic community, passed; however, immersion in a linguistic community did not guarantee passing for NSL signers and Unschooled Spanish Speakers. Nevertheless, the Spearman’s rho revealed a statistically significant inverse relationship between the amount of time an individual lacked a linguistic community and UC score ($r_s[15] = -.51, p = .05$).
However, no significant relationship between linguistic community and A/R FB was found ($r_{[15]} = .43, p = .11$). The Spearman’s rho also failed to reveal a statistically significant relationship between education and either FB task (UC: $r_{[15]} = .23, p = .42$; A/R: $r_{[15]} = .08, p = .78$).

In sum, as strength of linguistic community increased (with Homesigners having the weakest, and hearing Nicaraguans the strongest), the passing rate for A/R and UC increased. Notably, hearing Nicaraguans did not universally pass either task, in conflict with our prediction that membership in a linguistic community would enable passing. Also, the passing rates differed for A/R and UC, suggesting that these subtasks tap different aspects of FB (though typically developing American five-year olds pass both tasks).

Our confidence that participants, including the homesigners, understood the FB task derives from its design and from our years of experience working with these individuals on a variety of language and cognitive tasks. Unlike traditional FB tasks, the current task instructs via experience rather than verbally, and the task phases gave participants experiences prior to making predictions—the very experiences that a Simulation Theorist would expect to scaffold predictions about another’s choices.

The sticker phase provided the experience of making easy vs. hard (preference) choices, and then participants experienced earning or not earning incentives for correct predictions. An additional control is built into the UC phase: a correct prediction of matchbox could come about in three ways: the participant actually understood the other person’s FB; they forgot what was actually in the box, or they didn’t understand the task, and simply answered on the basis of the appearance of the box. However, in our dataset, whenever a participant did not succeed, they had predicted the lock/lockbox, evidence that they had not forgotten what was in the box and were actually trying to predict the tool match. Moreover, the homesigners never indicated that they thought they were being fooled. To ensure this trust, the items for the FB tasks were never moved from the participant’s sight.

**Study 2: Executive Function**

Executive function (EF) has been studied extensively (e.g., Anderson, 2002), and has been shown to relate to ToM development (e.g., Carlson et al., 2004). de Villiers & de Villiers (2012) summarizes the mixed findings regarding executive function’s role in A/R tasks vs. other FB tasks. Study 2 explores these relationships in the Nicaraguan groups, who vary in language and educational background. To our knowledge this study is the first to investigate EF abilities in these groups. We tested their Executive Function abilities using a simple inhibitory control (mixed congruent/incongruent) task (Shusterman et al., 2012) that required minimal verbal instruction and required only a behavioral response (vs. a verbal response). This task was chosen because it has been used with young deaf populations in the United States, and because the conflict of rules parallels a FB trial: one must suppress one rule (or truth) for another. This EF task also provided a measure of non-verbal intelligence with minimally verbal instructions – an important methodological consideration given the lack of standardized non-verbal IQ tasks that could be used across the three participant groups in our study.

**Method**

**Participants**

All participants were the same as in Study 1.

**Materials and Procedure**

The experimenter sat across from the participant, with two 6”-diameter light-buttons on the table between them. In the instruction portion, the experimenter donned a black or white glove on her right hand and gesturally or verbally instructed the participant that when she tapped a button with that hand, the participant was to tap the same button. The participant was allowed to practice with feedback. The experiment had three testing phases: 1) Right hand; 2) Left hand; 3) Both hands. In the first testing phase, the experimenter tapped the buttons with her right hand in a prescribed pattern (3x on the left, 3x on the right, 6 total).

In phase 2, a similar procedure was used with the left hand, but with the opposite glove color (i.e., if the right hand had had a black glove in phase 1, then the left had the white). Laterality of glove color was counterbalanced across participants. Importantly, the left hand “rule” was that whichever button was tapped, the participant had to tap the other button. After a few practice trials with feedback, the task began with six button taps, as in phase 1.

In the third phase both gloved hands (one black and one white), tapped the buttons. Both rules from phase 1 and phase 2 continued to apply: when the right hand tapped a button, the participant still had to tap the same button;
likewise, when the left hand tapped a button, the participant had to tap the *other* button. Note that the glove color/ hand did not change from the control phases. Phase 3 contained 12 trials, with four possible hand/button combinations. All participants, including the homesigners, clearly understood the rules; in fact, homesigners enjoyed the task so much that they afterward donned the gloves themselves to play with others as a diversion.

![Figure 2](image)

**Figure 2.** Individual scores for each participant on phase 3 of the Executive Function task (two rules combined). Circle color indicates FB performance (see legend).

**Results**

All participants scored well above chance (50%) on control phases (right & left hands alone), (M$_{NSL}$=96%, M$_{SSL}$=95%, M$_{US}$=86%), showing they understood the task. The phase of interest is the third phase, which required the participants to maintain two conflicting instructions – to either tap the *same* or the *other* button, depending on the hand being used/glove color. Typically developing hearing children in the US scored 86% on this task (M$_{age}$=57 mo.) (Shusterman, et al., 2012). We thus set the passing criterion at 80% for the adults; given this, no homesigner passed, 5 out of 6 NSL Signers passed, and 1 of 5 USSs passed (Figure 1, grey bars).

Figure 2 presents individual EF scores, with FB scores indicated by marker color (see legend). The Spearman’s rho revealed a statistically significant relationship between years of education and EF score ($r_s[15] = .62$, $p = .01$). Homesigners and USS scored in the same range on EF, with the exception of one USS who scored 100% – this participant is the only Spanish speaker who reported *some* education (3 y), whereas all other USSs reported none.

**Discussion**

The fact that NSL Signers, the only group with significant educational experiences, tended to score higher than the other two groups, suggests that education may play a role in developing EF abilities. Somewhat surprisingly, 4 out of 5 Spanish Speakers did not perform at levels achieved by typically developing American 5-year-olds, suggesting that a native language alone does not ensure success on this task. These results suggest the possibility of an interaction between language and education. However, our design cannot distinguish between membership in a language community combined with education and education alone. A *community language* (even an emerging one) combined with moderate$^2$ levels of education apparently enabled NSL Signers to pass at a higher rate than the other two groups. Notably, NSL Signers’ lack of a conventionalized language and the fact that their primary language was not even used in their classrooms did not prevent the majority of them from achieving criterion.

**General Discussion**

We conducted two studies investigating the effects of language and education on false belief understanding and executive function performance in three Nicaraguan groups: Homesigners, who do not participate in a linguistic community and have little schooling; Nicaraguan Sign Language Signers, who were among the initial creators of this emerging language, and Unschooling Spanish Speakers, members of a linguistic community, with sparse education.

Our results suggest three conclusions: 1) a linguistic community is necessary, but not sufficient, to support success on FB tasks; 2) education, more than language, relates to executive function; and 3) executive function ability does not relate to false belief performance.

These results support theories in which language may play a crucial role in the development of theory of mind. According to simulation theorists, having an experience of a false belief shortly before predicting the response of another individual holding the same false belief in the very same situation would eliminate the need for language, and indicate that life experience and introspection supports the understanding of others’ beliefs and actions. Our results showed that life experiences, even those experienced immediately prior to prediction, did not help adults without a linguistic community. Thus, Simulation (and related) theories (e.g., Gordon, 1986; Goldman, 1992) cannot explain the behavior of homesigners, who presumably reflect on their own experiences, yet remain unable to leverage that reflection into successful prediction of others’ behavior.

We found the executive function results particularly striking. NSL Signers outperformed both the Homesigners and Unschooling Spanish Speakers, both of whom had very little educational experience. However, these groups differed greatly in their linguistic profiles: despite having typical language backgrounds, only one Spanish Speaker passed executive function. As it turns out, he was the only participant who had had some schooling. Thus, we suggest that education, more than language, underpins congruent/non-congruent EF abilities. Ostrosky-Solis and colleagues (2004) found no differences between participants with and without education on a task (the

---

$^2$ The teachers of Cohort 1 signers used spoken Spanish and did not sign; thus, the reported education levels are likely overestimates relative to typical US levels.
“opposites” task) that was similar to our control Phase 2. We have not located any studies that examine language or education effects on a Phase 3-type task, which requires participants to manage both same- and opposite-responses. The orally-educated deaf children in Shusterman et al.’s (2012) study showed delays in executive function relative to their hearing counterparts. We speculate, based on our results, that later improvements in these deaf children’s EF would stem from additional educational experience rather than from further linguistic development.

Contrary to our hypothesis, the participant groups patterned differently on false belief and executive function, suggesting that EF ability does not drive FB performance.

We acknowledge some limitations. First, Homesigners, NSL Signers, and Unschool Spanish-speaking adults are extremely rare and difficult to recruit and test, yielding small sample sizes. Also, we did not administer any standard IQ tasks. Over many years of interaction with the research group, the homesigners have not exhibited signs of congenital cognitive deficits, and their performance on perceptual matching and mental rotation tasks is in the same range as that of their hearing family members. Future work will utilize implicit measures of FB performance, such as eye-gaze or looking time measures.

**Conclusion**

Our results suggest that language primarily contributes to success on false belief tasks, whereas education is related to executive function success. The homesigners’ inability to pass either type of task indicates that life experience alone isn’t enough to overcome the consequences of lack of participation in a linguistic community, lending support to theories that emphasize language in ToM development, rather than experiences (e.g., Simulation Theories). Some aspects of language require participation in a linguistic community to develop. While our results support a critical role for language, they do not address which aspects of language structure are essential. Further research is necessary to assess the specific relationships between participation in a linguistic community and education on executive function and theory of mind understandings.

**Acknowledgments**

We are grateful to the hearing and deaf participants for allowing us to learn from them. We thank Dr. Claudia Molina and Leybi Tinoco for fieldwork support in Nicaragua, Jennie Pyers for incredibly helpful discussions regarding the experiential false belief task, Devika Prasad for coding assistance, and Emily Carrigan and Russell Richie for comments. Funding provided by NSF IGERT, Graduate Research fellowship and a UConn Center for Latin American Studies/El Instituto grant to DG, and NIH P30 DC010751 to MC and D. Lillo-Martin.

**References**


