**Chelsea R Lide**

Honors Psychology: Mind, Brain and Behavior

Harvard Class of 2016

*Language, Goals and the Selective Learner:*

*How Syntax Guides Infants’ Interpretation of Event Structure*

**Writing Schedule**

**Jan 4** – Complete Experiment 1 write-up (methods, results, discussion)

**Jan 8** – Complete Experiment 2 write-up with skeleton data (methods, results)

**\*Jan 20** – Complete Abstract, Introduction and Discussion; First full draft of thesis

**Edits Log**

|  |  |  |
| --- | --- | --- |
| **Date** | **Section** | **Summary** |
|  |  |  |
| **Jan 3** | Experiment 1 (all) | Updated edited version of Experiment 1; added results and discussion |
| **Jan 4** | Exp 1: discussion; Exp 2: methods | Updated discussion section; started draft Experiment 2 Methods |
| **Jan 5** | Exp 2: methods | Updated Experiment 2 methods |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**Abstract**

**Introduction**

“The trouble is that an observer who notices *everything* can learn *nothing*”

—Lila Gleitman

Humans are innately communicative beings, but how is it that we come to develop our capacities for language? Among the first theories to be proposed was a simple association mechanism by which learners heard the unfamiliar word in the presence of its referent, giving rise to seemingly intuitive word-to-world mapping between language and concept (Locke, 1690). However, a number of conceptual truths confound this proposal. First is the pure statistical and computational power required to properly match an unfamiliar word to an infinitely long list of possible referents (Quine, 1960). This account, which would be a slow, laborious and error-ridden process, does not align with the observed speed and accuracy with which young children acquire language (Carey, 1978). Additionally, this mechanism does not speak to how we come to acquire words for concepts that are abstract, unobservable, dynamic, or ones that necessarily co-exist (e.g. cars and tires, or bunnies and floppy ears).

In this way, the challenge of acquiring a language is an induction problem: a finite source of information (our limited exposure to a word) yields an infinite number of hypotheses (all of the possible referents to which a word may refer). As Steven Pinker (1994) points out, induction problems are often overcome by the presence of natural constraints on these possible hypotheses. By systematically reducing the number of plausible referents in a conceptual space, word-to-world mapping becomes more robust and efficient. Indeed, constraints of this kind have been found within the domain of noun learning (see Markman, 1991; Markman & Hutchinson, 1984; Markman & Wachtel, 1988), but to what extent do these principles extend to the conceptually more difficult task of verb learning?

*Selective Attention: A Moderate Theory of Syntactic Bootstrapping*

Syntactic bootstrapping, a term coined by Lila Gleitman (1990) in her book on language acquisition, is the verb-learning analogue to noun-learning biases. The principle claim of syntactic bootstrapping is that children use the structure of the sentence in which a novel verb appears to bootstrap its meaning (Gleitman, 1990). This reduces the number of possible concepts that may be considered as a word’s referent, and also explains how verbs to describe abstract or unobservable concepts are acquired. Importantly, this “database” of potential concepts is categorical, rather than probabilistic, meaning that less evidence is required to reliably and accurately map a word to a concept. Theoretically and anecdotally, Gleitman’s proposal seems to robustly account for both the patterns and the nuances of children’s language acquisition. However, her argument is not without its limitations.

A number of qualifications to Gleitman’s theory have been suggested (see Pinker, 1994, for the most direct critique), but perhaps none as fundamental as those calling into question the actual power and purpose of the subcategorization frames. Pinker (1994) argues that these syntactic structures are not actually encoding the true *meaning* of a word, as stronger interpretations of the syntactic bootstrapping theory would suggest. Instead, they are acting as a sort of “zoom lens” that helps the listener discern to what aspect or perspective of the event structure a verb is referring. Pinker believes, and Gleitman concedes, that some information is just not available in subcategorization frames, and needs to be gleaned through observation. Thus, if we wish to make sense of how subcategorization frames and observation systematically facilitate the acquisition of novel verbs, we must turn to research that explores to what aspects of event structure young language-learners are able to attend.

*Understanding Events in Infancy: Principles of Rationality*

Infants’ comprehension of the world around them is far from the “blooming, buzzing confusion” originally suggested by famed psychologist, William James (1890, pg. 488). Remarkably, even very young infants demonstrate adherence to a highly systematic schema of event interpretation. Amongst the most widespread and foundational components of this schema is their comprehension of—or, as some would suggest, obsession with—goals.

In his seminal paper, Andrew Meltzoff demonstrated that 18-month-olds would perform the *intended* action of an experimenter, both in conditions where the experimenter’s attempts were successful *and* when they were not (1995). This work demonstrated infants’ ability to infer the intentions of the adult experimenter, with focus given to the goal (i.e. successful outcome) of the action, rather than the means used, in vain or not, to achieve it. Further, Woodward (1998) found this same goal-fixation behavior in infants as young as six months. By this age, she claims, infants have a “mature understanding of goal-directed action,” and this helps them to attend to the most relevant aspects of an event (1998, pg. 1).

In light of this, the contemporary belief was that infants are acutely able to distinguish between goals and means (Meltzoff, 1998). However, unlike the focus given to how infants interpret goals, a lack of exploratory work left researchers with a markedly impoverished understanding of how infants conceptualize means. In fact, research at the time suggested a highly simplistic view: when infants were given the opportunity to imitate an event, they would exclusively imitate the means of the demonstrating actor with little variation or room for interpretation (Meltzoff, 1998). It wasn’t until Gergley and his colleagues (2002) published subsequent work on infants’ understanding of means that the true complexity of infants’ comprehension of goal-directed events was revealed.

In what came to be known as Gergley et al.’s (2002) rational actor imitation paradigm, 14-month-olds saw an experimenter achieve a goal (turning on a light box) by utilizing a novel means (leaning forward to touch her forehead against the light box). While all infants saw the same novel manner produce the same goal, the context of the events varied between the two conditions. In the *hands occupied* condition, the experimenter had her hands concealed underneath a blanket, which she was holding tightly around her shoulders. In the *hands free* condition, her hands were placed on the table to either side of the light box, clearly visible to the child, while the blanket was loosely draped around her shoulders.

What Gergley and his colleagues found was that imitation responses patterned differently between the two conditions—a surprising result given Meltzoff’s and others’ impoverished understanding of infants’ expectations about means. Infants in the *hands occupied* condition opted to use their hands more often than their heads when imitating the novel event. The researchers suggested that the infants were attributing the experimenter’s novel behavior to the fact that she had her hands unavailable to her, as they were holding the blanket around her shoulders, and thus had no other means but her head to achieve the goal. Given that the infants did, in fact, have use of their hands, they were more likely to make use of that more practical manner. In contrast, infants who saw the experimenter perform the novel head touch, even when her hands were clearly available to her, more closely imitated the novel means demonstrated. They were supposedly reasoning that, if the experimenter had the option to use her hands yet chose to use her head, there must be something important or essential about this novel choice, and thus, that proper imitation necessitates it (Gergley et al., 2002).

Infants are not only demonstrating expectations about goal-directedness, but about the means to achieve those goals by as early as 12-months-old (Schwier et al., 2006). When combined with previous theories on infants’ event comprehension, findings from rational actor imitation paradigms give rise to a more robust and accurate depiction of how infants interpret events as they unfold. First, it is known with relative certainty that infants expect humans (but not necessarily inanimate actors) to be goal-directed. Infants are also capable of deducing what means are most efficient to achieve the goal, and importantly, utilize that particular context’s unique constraints to inform their judgment. Lastly, once infants have developed a naïve theory about the most efficient means to achieve a goal in a specific context, they then expect actors around them to perform in accordance with this efficiency. These characteristics were confirmed in a study done by Phillips and Wellman (2005), who found that 12-month-olds were only surprised to see an experimenter use an indirect arm reach to grab a ball when the context made the manner novel—that is, when the experimenter could have used a direct reach, instead. When the situation was changed and a barrier was placed between the experimenter and the ball, the novel arm arc was no longer interesting to the infants, given their appraisal of the new context’s constraints (Phillips & Wellman, 2005).

Taken together, this evidence suggests that infants are able to systematically parse dynamic events into their component features, and further, that they are then able to use these complex understandings to inform both their judgments of present events, and their expectations about subsequent actions. The ability to attend to the various components of an event structure is analogous to the “zoom lens” mechanism suggested by a moderate interpretation of Gleitman’s (1990) theory of syntactic bootstrapping. In particular, the emergence of the goal-means distinction early in the development of event perception could suggest a similar mechanism within the realm of language. In combination, this would provide both the subcategorization frame and extralinguistic (i.e. observational) information infants require to constrain the infinite list of possible concepts to which a novel verb may refer.

*Linguistic Manifestations of the Goals-Means Distinction*

In linguistics, there are a number of dimensions upon which verbs may be classified. One of these distinctions divides an event structure on the basis of either goals or means, similarly to the relevant research on infants’ event perception. In what’s known as the manner versus outcome distinction, single verbs, as well as entire syntactic frames, encode information about either *how* something is done (in the case of manner verbs/phrases) or the *result* of the thing done (in the case of outcome verbs/phrases). On the word-level, take the following sentences:

1. Jane *wiped* the dirty table
2. Jane *cleaned* the dirty table

From sentence (1), we may infer what means are occurring during the event—that is, that there is “wiping” taking place—but there is nothing encoded in the verb itself that reveals the outcome of the event. This is to say that both of the following sentences are perfectly sensible:

1. Jane wiped the dirty table, *and then it was clean*
2. Jane wiped the dirty table, *but there was still food everywhere*

In both cases the same means are used, yet result in very different outcomes. Compare this to the contrasting example of an outcome verb in sentence (2). Here, we can infer the end result that has occurred—namely, that the table is now clean—but we are none the wiser to how this goal was achieved; nothing in the verb’s meaning encodes whether the means involved using soap or a sponge or a power washer, etc.

The same distinction also applies on the structural level. Take the following subcategorization frames:

(5) Mark \_\_\_\_\_\_\_ *to* his neighbor

(6) Mark \_\_\_\_\_\_\_ his neighbor

Verbs that fit sensibly into the syntax of sentence (5) encode manners (e.g. *yelled, waved, sang*), while logical completions of sentence (6) encode outcomes (e.g. *hit, hugged, saw*). What becomes clear in these examples is the constraining power of the manner versus outcome syntactic distinction as a subcategorization frame. Importantly, manner and outcome verbs—and by extension, the syntactic frames into which they fit—are nearly always in complementary distribution (Rappaport Hovav & Levin, 2008). This added restriction means that this contrast is poised to be a particularly rich and reliable cue. What then follows is to explore whether children are sensitive to this cue in the same way that adults are, and if so, whether this linguistic information can modulate, or shift the “zoom lens” on their interpretation of extralinguistic events. According to Kline and Snedeker (2015), both of these claims are true by the time children reach two years of age.

In their study, all participants saw the experimenter perform a novel head touch in the *hands occupied* position, identical to the one featured in Gergley et al.’s (2002) rational actor imitation paradigm. This establishes a baseline response—namely, that without any other intervention, children should perform the more rational hand touch response when asked to imitate the novel event. The participants were then split into two language conditions: *manner* and *outcome*. In the *outcome* condition, children heard a novel verb within a goals-encoding syntactic frame (i.e. “I’m gonna *dax* my toy”). In the *manner* condition, children heard the same novel verb used within the context of a means-encoding sentence (i.e. “I’m gonna *dax to* my toy”). The experimenter performed the novel action, and then used the novel sentence to describe what she had just done. After repeating this demonstration process twice, the experimenter prompted the children to *dax*/*dax to* the toy (dependent upon condition). The measure of interest was children’s first contact with the toy.

When imitating the novel event, participants in the *outcome* condition performed proportionately more hand touches. This result is unsurprising, given that they received both syntactic and observational cues encoding the goal as the most important feature, thus privileging the more efficient means. In contrast, children in the *manner* condition performed the novel head touch more often, suggesting that the syntactic cue was salient enough to shift participants’ perspective toward the means, and pull them away from the more rational baseline response (Kline & Snedeker, 2015). 2-year-olds’ selective imitation of the feature of the event encoded in the syntactic frame used by the experimenter, even when this imitation was not the most rational means possible within the event’s context, demonstrates the use of the linguistic frame as a constraining mechanism used to guide attention in event structure.

The present research attempts to extend this work to 18-month-olds. As we have seen, infants at this age have a rich understanding of rational action, but can these expectations be modulated by language in patterns similar those found in 2-year-olds? Exploring this would reveal two principle insights: first, to what extent young infants are sensitive to the manner versus outcome distinction; and second, if infants are capable of using the manner versus outcome distinction as a cue—insofar as it creates a particular subcategorization frame—to shift their attention to the privileged feature (either means or goal) in their conceptualization of the event structure. From this, we may begin to chart out a developmental trajectory for this particular cognitive mechanism, and broaden our understanding of what subcategorization frames are useful to language-learners, when they become accessible to infants, and how exactly they facilitate word learning.

**Experiment 1**

Short intro to experiment needed?

**Method**

**Participants**

Participants were twenty 17- to 19-month-olds (range 17;00 to 19;00; 11 girls). An additional seven infants were tested but not included in the final analyses due to refusal to interact with the toy at test (*n* = 5) or experimenter error (*n* = 2). All children were recruited from a university database of interested families in the Cambridge area, and received a small toy and five dollars of travel compensation for participating.

**Materials**

The novel toy presented during the critical trial was a 12in x 4in x 10in box covered in green felt. On one side of the box surface was a large silver dome, which was meant to focus children’s imitation response to one local area, given their familiarity with acting upon buttons. A globe was situated a few inches away from the button on the box surface, and contained lights that would illuminate and spin upon activation (*figure X*). The globe’s handle was concealed within the box, and was wired to a button to facilitate hands-free operation of the spinning lights. A small camcorder was positioned facing directly perpendicular to the child to record their interaction with the toy.

**Procedure**

Families were greeted upon arrival to the lab, where the experimenter engaged the child in interactive free play in the lobby. At this time, parents were given instructions on how to neutrally respond to their children during the exploration period, where the participants were allowed to freely interact with the toy. To reduce any potential for biases, parents were also asked to refrain from giving explicit guidance on how to operate the toy, and told instead to give vague feedback such as, “hmm…I don’t know!” or “what do you think?” Parents were also informed that they should avoid specifically directing their child’s attention to the toy, as a lack of interest would be an equally meaningful measure of engagement.

Then, when the child appeared to be adequately comfortable socializing with the researcher, the family was escorted to a second room to begin the experiment. The testing room was a well-lit space that was empty except for two chairs, a table, and a curtain lining one of the sidewalls. The infant was placed in the parent’s lap and the pair sat directly across from the experimenter with the small table located in between them.

The study then began with a series of simple warm-up trials similar to the game Simon Says, in which a puppet, manipulated by the experimenter, would perform a simple action (e.g. clapping) and then encourage the child to imitate the action as well. This activity was geared toward preparing the infants to engage in imitative play. At the end of the warm-up trials, the puppet was put away and the novel toy was introduced.

In both the hands-exposed (*free*) and hands-occupied (*occupied*) conditions, the experimenter exclaimed that she was cold, and proceeded to wrap herself up in a blanket made of blue fleece. In the *occupied* condition, the experimenter used one hand to hold the blanket tightly around her shoulders, and the other to surreptitiously operate the toy out of view of the infant via the wired button. In the *free* condition, the experimenter loosely draped the blanket over her shoulders, and while doing so, covertly attached the button to a small piece of Velcro located on the underside of the table. The experimenter could then operate the toy by simply raising her knees to compress the button against the table. Once the button was secured, she placed her hands palms-down to either side of the toy, clearly visible to the child.

At this point, the critical sentence was introduced: either “Look! I’m going to *dax to* my toy!” in the *manner* condition or, “Look! I’m going to *dax* my toy!” in the *outcome* condition. This sentence was then followed by a demonstration the novel event, which consisted of the experimenter leaning forward to touch the silver button on top of the toy with her head while simultaneously activating the toy’s lights. This created the illusion that physical contact with the silver button caused the lights inside the globe to turn on and spin (see figure X, grid of what conditions looked like).

After performing the action, the experimenter repeated the critical sentence to describe the event that occurred (e.g. “Look! I *daxed(to)* my toy!). This procedure was then repeated a second time, such that by the end of the demonstration period, the child heard the critical sentence a total of four times. The sentence was introduced one final time when the toy was placed within the child’s reach, and the experimenter prompted the child to *dax(to)* the toy. When the child made first contact with the toy, the experimenter did not activate the globe as in the demonstration, but instead simply responded with the neutral, yet enthusiastic reply, “Okay! Now you can play.” This then initiated the exploration period.

During the exploration period, the experimenter told the children “they could play” or that “it was their turn” before walking off to another corner of the room and shuffling papers to look preoccupied. The main purpose of this portion of the study was to investigate children’s persistence when the toy failed to operate. After 60 seconds, or sooner if the child had begun to fuss, the experimenter returned to the table and encouraged the child to make one more attempt at contact with the toy, which was rewarded by activation of the lights and very enthusiastic praise. Children were allowed to play with the now-functioning toy for a little while longer before the session was ended, and the families were debriefed and thanked.

**Coding**

All sessions were videotaped in order to accurately assess each child’s first contact with the toy. Videos were viewed by the experimenter directly after the session, and coded as “hand touch,” “head touch,” “N/A” (for no response) or “fuss out,” when a child was unable to complete the experiment. Any contact made exclusively by a hand (i.e. pressing with palm) or finger (i.e. poking) was coded as a hand-touch. For the purposes of this particular paradigm, a head touch was inclusive of lips, cheeks, chins, and the like, in addition to the more straightforward forehead contact. Given 17- to 19-month-olds’ limited motor coordination, head touches preceded by the use of the hands as a helping agent (e.g. lifting the toy to their head) were considered valid head touches. Anecdotal evidence suggested that manifestation of children’s imitation attempts at this age are extremely varied, thus the more inclusive criteria for a head touch was designed to capture whether the children recognized the use of a novel body part, and sought to imitate that novelty, as well.

**Results**

To facilitate a more robust analysis of the data, a 2 x 2 (hands x language) between-subjects design was used. Running all four possible conditions simultaneously allowed for various patterns in the data to be interpreted more holistically. For example, if participants in the hands exposed conditions performed more head touches than those in the hands occupied conditions, regardless of language condition, this would support a successful replication of Gergley et al.’s (2002) finding that \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. Further, if participants in the manner conditions performed more head touches than those in the outcome conditions, regardless of hand condition, this would support two conclusions: first, that the semantic difference between manner versus outcome is reflected syntactically, and is salient to infants in our target age range; and second, that this difference is a cue that maps to the interpretation of events as they occur in the world (and as they are perceived by the child).

Taken together, these predictions give rise to a sort of cue gradient, in which infants in the exposed/manner condition are given the most head-touch eliciting cues, and are thus pulled farthest from the hand-touch baseline response. Infants in the occupied/outcome condition would then be situated on the opposite end of the spectrum, receiving the fewest head-touch eliciting cues, and consequently performing more hand-touches. Examining the pattern of results for infants in the two intermediary conditions, who received both head- and hand-touch eliciting cues, presents the opportunity to measure cue “strength” or “dominance” when conflicting cues are conveyed simultaneously.

The frequency of each first response type, either head touch or hand touch, is presented in Table X. Not only does this pattern of results markedly diverge from the proportional differences we would expect to see if our hypothesis were true, it clearly shows a failure of all but two participants to perform a head touch at any point during the trials. Contrary to our predictions, nearly 90% of infants performed the head-touch baseline response, regardless of condition.

Should I go on (& mention gender/mcdi differences per trial & log regression), or futile at this point…?

This null result is further proven through logistic regression, which shows no significant relationship between condition and predicted first response, STATS.

While each trial included an exploration period in which infants were allowed to freely engage with the novel toy, their behaviors were not coded as a part of this analysis. Any patterns therein would not be interpretable within the scope of this particular paradigm because of its failure to replicate Gergley et al.’s (2002) results. Without producing the expected differences between the hands conditions, where those who saw hands exposed performed more head touches, we cannot reasonably draw conclusions regarding the effects of language, our true variable of interest.

**Discussion**

Experiment 1 was designed to test the relationship between the structure in which a novel verb is presented, and the interpreted meaning of that verb. If it were the case that children used the semantic cues, conveyed through syntax, to guide their interpretation of a novel event, then we would have expected to see a cue gradient between the conditions. Participants who heard and saw cues that suggested the manner to be the essential feature of the novel event should consequently seek to more closely imitate this feature, and thus perform the most head-touches. Participants in the outcome-focused condition should conversely prefer to take the simplest means to achieve the outcome: the more efficient hand-touch.

However, we instead found that nearly all participants in Experiment 1 performed the baseline response of a hand-touch, regardless of condition. Given our inability to replicate Gergley et al.’s (2002) paradigm, valid conclusions cannot be drawn regarding infants’ ability to perceive the manner versus outcome distinction, or to use this distinction to guide their verb learning. Thus, the question of the effect of linguistic structure on children’s interpretation of verb meaning remains unresolved. If we wish to more thoroughly explore the relationship between the manner versus outcome distinction and how infants use such cues to guide their verb learning, modifications to our paradigm must be made.

The first consideration made was in regards to potential physical constraints within the paradigm. Perhaps it is the case that infants in this age range are physically unable to perform the motor functions necessary to complete a head-touch. This action requires core strength, upper body strength, and the skilled coordination of the two. However, previous research, including Gergley et al.’s (2002) study featured infants two to four months younger than our age range successfully completing head-touches (*for reference, see* \_\_\_\_\_\_\_\_\_\_\_).

Our focus then shifted to the toy, itself. After thorough review of the previous studies that used any remotely similar type of light box stimulus within the context of a rational actor imitation paradigm, two alterations became evidently necessary. First, were the actual dimensions of the toy. Most toys used in the other studies ranged in height from X to Yin, markedly shorter than our 10in tall toy. Lowering the toy not only puts it within a physical range that is more comfortable for the infants, but also makes the novel head-touch action far more salient, by requiring a full bend at the waist by the experimenter, rather than the simple head tilt required to reach the taller toy.

The second change to the toy streamlined the perceived relationship between the novel head-touch and the activation of the toy. To accurately recognize the novel event using the original stimulus, the infant must necessarily understand a relatively complex causal model. They must interpret that the experimenter’s acting on one side of the toy (the button) causes an effect in a visually distinct entity located on the opposite side (the globe). Thus, in Experiment 2, the button was removed, and the experimenter instead acted directly upon the globe, which was centered in the toy. This modification reduces the complexity of the action, while still providing a focused location toward which infants may direct their imitation response.

Using the original 2 x 2 design, in conjunction with the updated toy, Experiment 2 was an attempt to replicate Gergley et al.’s (2002) original head touch patterns. If successful, we would then be situated to explore the added effects, if any, of syntax on infants’ interpretation of the meaning of a novel verb.

**Experiment 2**

**Method**

**Participants**

Participants were X 17- to 19-month-olds (range 17;05 to 19;00; 17 girls). An additional four infants were tested but not included in the final analyses due to inability to complete the experiment (*n* = 3) or parental interference (*n* = 1). All children were recruited from a university database of interested families in the Cambridge area, and received a small toy and five dollars of travel compensation for participating.

**Materials**

The novel toy presented during the critical trail was a 12in x 10in x 3in box covered in green felt. The globe and its handle were laid flat within the shallow box, such that the globe was partially protruding from the box’s surface (*figure X*). This created the illusion that the toy was a fuzzy green box with a button-like half-dome of lights situated in the center. A thin yellow ring made of construction paper was placed around the circumference of the globe, both to draw the child’s attention to the globe and to make it more visually appealing. This toy was wired similarly to the one originally used in Experiment 1 in that the handle was connected to insulated wires leading to a button that could be operated in both the hands free and hands occupied conditions. A small camcorder was positioned facing directly perpendicular to the child to record their interaction with the toy.

**Procedure**

The procedure used in Experiment 2 was identical to that used in Experiment 1, with only minor changes made to the testing room and warm-up period. Following the procedure of previous experiments that used light boxes within the context of rational actor imitation paradigms (e.g. \_\_\_\_\_\_\_\_\_\_\_\_), the lights in the testing room were slightly dimmed to increase the salience of the toy’s activation.

Additional toys, including a number of small stuffed animals, some building blocks, and a textured red ball, were displayed on the table at the start of the experiment. The experimenter engaged in free play with the child using these toys as a means to acclimate them to the unfamiliar room, as well as to further familiarize them with the experimenter. Toward the end of the free play, attempts were made to have the child interact with the experimenter directly, by passing the ball back and forth across the table. This passing game was also a seamless way to introduce the puppet, who then began passing the ball to the child. The ball was put away once the experimenter felt that the child was sufficiently familiar with both her and the puppet, at which point the Simon-Says-like warm-up task from Experiment 1 was initiated.

Unlike in the first experiment, parents were instructed to participate in the warm-up game if their child was particularly shy or reluctant to engage. This was done, for example, by saying, “Hmm, I think mommy/daddy knows how to clap. Let’s all clap! Look! We’re all clapping!” [This is mostly anecdotal….do I need to prove it statistically? Is it even worth mentioning?]

When combined, the extended and more imitation-focused warm-up period resulted in a lower proportion of participants fussing out (6%) or failing to interact with the toy at test (10%) as compared to Experiment 1 (x/27 and 19%, respectively).

As in the first experiment, the puppet was put away at the end of the warm-up trials. The novel toy was then introduced, demonstrated, and explored by the child, who received enthusiastic praise before the session was ended. Families were then debriefed and thanked.

**Coding**

All sessions were videotaped in order to accurately assess each child’s interactions with the toy. Unlike in Experiment 1, each individual session of Experiment 2 was coded in two iterations: (1) to record the participant’s first contact with the toy; and (2) to chronicle the full range of actions performed during the exploration period.

**First Contact*.*** To indicate participants’ first contact with the toy, the experimenter reviewed videos directly after each session. Using the same criteria as in Experiment 1, each first response was designated as a “hand-touch,” “head-touch,” “N/A” (for no response), or “fuss out” when a child was unable to complete the experiment.

**Exploration Period*.*** At a later date, the experimenter used the video annotation tool VCode to catalogue behavioral data during the exploration period. Each video was coded according to three clusters of interest: duration, engagement and action.

***Duration.*** The duration of several key events in each session was recorded, with onset and offset accuracy on the scale of milliseconds. These events included the length of the total trial, warm-up period, demonstration length, first response window and exploration period (see Table X for specific coding criteria).

***Engagement.***  Children’s engagement during the exploration period was coded on two levels. The lower level, simply titled “Engagement,” was meant to capture the coarsest measure of time the child spent interacting with the toy. Because this metric was designed to be broad, it characterized any combination of (1) looking at the toy (at least at the start); (2) touching the toy; (3) discussing the toy in some capacity (e.g. talking about it directly, asking for parent’s help to fix it, etc.)

Moving up hierarchically, the next level of engagement describes “Intentional Body Actions (IBAs).” These instances still only capture very broad behavioral patterns, but narrow in on the types of physical actions of particular interest to our hypothesis. IBAs include physical contact with the toy, which is not exclusive to the globe, and looking at the toy *at the start* (but not necessarily for all) of the contact period.

***Action.*** Events within the category of *Actions* are the most detailed account of behaviors performed during the exploration period. They also offer the most fine-grained analyses of trends or correlations between specific behaviors or behavioral patters and predictors such as a participant’s condition or the length of their warm-up period, for example. The actions coded included the first action (first contact), head-touches and hand-touches (see Table X for specific coding criteria).

When coding for head- and hand-touches, the disparity between the efforts necessary to complete each action became abundantly clear. Performing a single head-touch is a multi-second endeavor; the same amount of time in which a multitude of hand-touches could have been performed. The rapid banging, for example, that could be done is nearly impossible to replicate using one’s head, and out coding scheme needed to accurately account for this incongruence. To do so, head-touch instances were recorded at face value: when a child leaned forward toward—or raised—the toy to touch it against their face/head. Hand-touches, however, were classified using a stricter, time-lapse criterion. For any hand contact to be rightfully considered a hand-touch, there must have been a half-second delay between its onset and the offset of any hand contact that preceded it. Doing this reduced any statistical noise created by repeated hitting, and is also preferable to the alternative method of simply measuring the length of time spent engaging in each respective action, which does not sufficiently account for the inherent differences between the response types.

**Results**

**Discussion**

**General Discussion**

**Conclusion**