Language, Goals and the Selective Learner:

How Syntax Guides Infants’ Interpretation of Goal-Directed Events

An Honors Thesis presented

by

Chelsea Rose Lide

to

The Department of Psychology

in Partial Fulfillment of the Requirements

for the Degree of Bachelor of Arts

with Honors in the Subject of Psychology

(Mind/Brain/Behavior Track in Cognitive Science)

Harvard College

Cambridge, Massachusetts

March 10, 2016

***Acknowledgements Here***

Jesse, Susan, Melissa, Danielle

LDS Grant Info

Mellon Funding

Mom

**Table of Contents**

1. Abstract…......……………………………………………………………………..…..…00
2. Introduction………………………………………………………………………..…..…00
   1. Heading...............................................................................................……......….00
   2. Heading...............................................................................................……......….00
   3. Heading...............................................................................................……......….00
3. Experiment 1………..…………………………………………………………..……..…00
   1. Methods……………………………………………………………………….….00
   2. Results………………………………………………..……………………….….00
   3. Discussion…………………………………………………………………….….00
4. Experiment 2………..……………………………………………………………..…..…00
   1. Methods…………………………………………………………………….…….00
   2. Results………………………………………………..……………………….….00
   3. Discussion…………………………………………………………………….….00
5. General Discussion…………………………….…………………………………..…….00
6. Conclusion……………………………………………………………...….……......…...00
7. References……….……………………………………………………...….……......…...00
8. Appendix.......…………………………………………………………...….……......…...00
   1. Descriptive Analyses Figures: Experiment 1.........................................................00
   2. Descriptive Analyses Figures: Experiment 2.........................................................00
   3. R Code...................................................................................................................00

**Abstract**

Studies have shown that children are able to use syntactic information about the structure in which a verb appears to guide their verb learning. At the same point in development, children also have access to robust schemas that allow them to interpret goal-directed events on the basis of an agent’s perceived intentions and rationality. Previous research has demonstrated that 2-year-olds integrate this structural and behavioral information to inform their interpretation of a novel verb (Kline & Snedeker, 2015). The present study extends this paradigm to 18-month-old infants in an attempt to outlinethe developmental trajectory of this ability. Twenty infants (mean age = 17;21, SD = 18 days, 11 girls) participated in Experiment 1. However, the results suggested that paradigm limitations unrepresentatively constrained participants’ imitation responses. Thus, the paradigm was modified, and forty-five infants (mean age = 17;27, SD = 17 days, 20 girls) participated in Experiment 2. Despite successfully eliciting the full range of possible imitation responses, the data suggest that these behaviors were not reliably predicted by participants’ conditions. This null result is discussed within the context of 18-month-olds’ linguistic and cognitive capacities. Particular focus is given to the probable influence of their underdeveloped theory of mind, which suggests constraints on the domains in which the rational actor imitation paradigm may be used as a valid measure of infant cognition.

**Keywords**: *syntactic bootstrapping, novel verb learning, rational imitation, theory of mind*

“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? By studying children’s linguistic abilities, we are able to characterize the emergence and development of this faculty. During this same period when children are beginning the process of language acquisition, they are simultaneously forming rich cognitive schemas about their social and physical worlds. Probing interactions between cognitive and linguistic development thus presents an opportunity to establish a mechanism for understanding the nuances of either individual capacity through the lens of the other. This in turn leads to a more robust characterization of the formation of adult cognition and language.

However, the basis of these claims rest entirely upon the assumption that language and cognition exist in relation to one another beyond any simple maturational coincidence. If we wish to study the nature of the interaction between these domains of knowledge, we must first establish that there is a connection between them. This begins by evaluating why it would even be useful for language and cognition to operate in tandem in the first place.

A possible point of convergence originates from the observation that language acquisition is an extremely challenging conceptual task. Unlike the grammar and vocabulary lessons children receive in school later in life, the most systematic and fundamental building blocks of language are not explicitly taught to infants (NAME, 0000). In addition to this lack of explicit teaching, children must formulate their theories of language without the presence of negative evidence (Marcus, 1993). While the usefulness of having access to the grammatical possibilities of a language still remain, this leaves any language production mechanism unaware of what utterances are *not* grammatical possibilities, and thus underpowered. This is further exemplified by the fact that in natural language discourse, parents rarely produce ungrammatical utterances or outwardly correct the incorrect grammar of their infants (Marcus, 1993).

Even more fundamentally, language acquisition is a complex challenge by virtue of being an induction problem: our limited exposure to the world (a finite source of information) yields an infinite number of possible “hypotheses” (all of the possible referents to which a word may refer). To fully grasp the meaning of a word, children must successfully pair an utterance to a referent that may be one of many features in a conceptual space, or simply not present at all. Parsing these pairings requires immense computational power, and simple association theories (cf. Locke, 1690) fail to explain even more conceptually complex tasks, such as how children acquire words for necessarily coexisting terms (e.g. “car” and “tire”) or subsets (e.g. “animal,” “dog,” “Fido”).

Yet in spite of this, children nevertheless acquire language rapidly and with far fewer errors than would be expected in light of these challenges (Carey, 1978). This would suggest that other sources of information inform and guide the word-learning process. Indeed, previous studies have shown that children utilize a number of domains of knowledge in language development, including social expectations about shared eye gaze toward the referent (NAME, 0000), and physical knowledge of relationships between objects in the world (NAME, 0000).

There also appear to be a number of innate constrains that facilitate language acquisition. For example, children are more likely to assign a novel label to a noun that does not already have a name (NAME, 0000), and will ascribe that label to denote the entire object, rather than an individual feature of that object (NAME, 0000). Constraints of this nature reduce the number of plausible referents in a conceptual space, lessening the statistical complexity of properly matching words to their real-world referents. When combined with the cognitive expectations developed through other domains of knowledge, a clearer model of the functional relationship between language and cognition begins to emerge. One domain in particular where this schema may be applicable is in the realm of verb learning.

Previous literature suggests that, at the same point in development, children are able to draw information from the syntactic structure in which a verb appears (Gleitman, 1990) while simultaneously forming complex expectations about agents and their actions (e.g. Gergley, 2002). At the surface, there is an intuitive benefit in concurrently learning about actions and the verbs meant to describe them. What remains to be explained is how these distinct sources of information interact to jointly shape children’s holistic interpretation of events as they are described.

To explore this question further, we will examine both how children derive semantic information from syntax, and what exactly the nature of children’s expectations about agents/actions are. This will inform how each in turn may promote the acquisition of verbs. Finally, original research will be presented, contributing to our current understanding of the developmental origins of the link between language and cognition.

**Syntax as a Source of Meaning**

To overcome the challenge of word learning, children are able to use the syntactic frame in which a verb appears to help deduce its meaning. This process, known as syntactic bootstrapping (Gleitman, 1990), is grounded in children’s understanding of the mapping rules between semantics and syntax. The function of these rules can be illustrated in the differences between the following sentences:

1. Jane sneezed
2. Jane hugged Sarah
3. Jane put the book on the shelf

The structure of each sentence follows from the semantic content of the verb contained therein. For example, the meaning of the verb *hug* in sentence (2) implicates that there are two arguments involved: the “hugger” and the “hugee” (Jane and Sarah, respectively). Syntactic bootstrapping suggests then that attending to syntactic cues, such as the number or order of arguments, helps children reduce the amount of possible meanings they will consider attributing to a novel verb. Thus, if a child is presented with a novel verb that takes three arguments, they infer that it is more likely to be a verb of transfer, for example (e.g. similar to the verb *put*, which necessitates (1) a “putter”, (2) a thing put, and (3) a location placed). Relatedly, other classes of verbs are simultaneously ruled out but virtue of not satisfying the expected structural form (e.g. the perception verb *look*, which takes only two arguments: the “looker” and the thing seen).

The type of information conveyed through syntax also extends beyond the simple counting of arguments. For example, two-year-olds are attuned to the transitivity of an utterance, which allows them to form expectations about agents *causing* actions, rather than simply engaging in them (NAME, XXXX). This remains true even when the number of agents is held constant: when children hear the novel transitive sentence, “Big Bird *is gorping* Cookie Monster,” they look longer at the scene in which Big Bird is causing Cookie Monster’s action than the scene where the two characters are performing the same action independently of one another (Hirsh-Pasek, Gleitman, Gleitman, Golinkoff, Naigles, 1988). The use of structural information as a conduit for meaning is evident in infants by as early as 9-months-old (Yuan, Fisher, Snedeker, 2012), and encodes information ranging from transitivity (NAME, XXXX) to causality (NAME, XXXX) to path and manner of motion (NAME, XXXX).

In addition to these cues, a lesser-studied source of verb-relevant information exists in what is known as the manner versus outcome distinction. This distinction encodes two forms of information: *how* something is done (manner) or the *result* of the thing done (outcome). These features represent important aspects of an event an observer may attend to, or that a speaker could wish to emphasize. At the lexical level, this information can be conveyed solely through the semantic contents of a verb. For example, take the following pair of sentences:

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

From sentence (4), 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 (i.e. *wiping*), yet result in very different outcomes. Compare this to the contrasting example of the outcome verb, *clean*, in sentence (5). After hearing the sentence, “Jane *cleaned* the dirty table,” we are able to 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 semantic contents of a verb also influence the broader syntactic structure the verb is likely to appear in. This is directly analogous to the way that a verb, like *put*, can probabilistically specify the number of arguments a sentence that contains it takes. For example, manner verbs, like *descend* or *dance*, will likely take a prepositional and/or an adverbial argument by virtue of encoding information regarding to location and/or means. Thus, if presented with a novel verb in a syntactic frame that contains a prepositional phrase (e.g. “Mark *glorped* to his neighbor”), it is more likely that the unfamiliar verb will be semantically related to other lexical items of manner (e.g. *waved*, *called*) than those encoding outcomes (e.g. *melted*, *exploded*).

Importantly, these inferences are markedly probabilistic. There are many instances where a verb typically considered to encode an outcome, may operate within an argument structure also semantically viable to manner verbs (e.g. “She *broke* into the kitchen”). This is to say that the structure in which a verb appears does not strictly govern its semantic contents, nor vice versa. This also highlights an important qualification of Gleitman’s theory of syntactic bootstrapping. While there is strong linguistic evidence and a number of experimental studies that suggest that correct links may be drawn between syntax and semantics, the derivational power of this relationship must not be overextended.

A limiting constraint arises from the fact that there are far fewer distinct clausal structures than possible verb meanings (Gleitman, 1990). Surely syntactic cues that bias a listener’s interpretation of a novel verb to any one of the possible semantic categories (e.g. *cause*, *transfer*, *cognition*, *perception*, etc.) may serve as a useful constraint, but these categories are ultimately too coarse to convey the full semantic contents of a verb. In this way, Gleitman’s (1990) metaphor of syntactic frames as a “zoom lens” is an apt characterization of their functional scope: syntax may guide attention to the aspect of an event a verb refers to, but as Pinker notes:

[N]o amount of lens fiddling can fix the vastly greater number of degrees of freedom defined by the potential contents of the picture – whether the lens is aimed at a still life, a nude, a ’57 Chevy, or one’s family standing in front of the Grand Canyon (Pinker, 1990: 399)

To supplement the limited, though useful, information syntax can provide about meaning, non-linguistic cues may be recruited to help a listener more comprehensively deduce the full semantic contents of an unfamiliar verb. This represents another instance where the challenge of language learning is facilitated by cognitive mechanisms in other domains. Specifically, the meaning of a verb meaning may be constrained—and thus, more easily deduced—by utilizing cues from the manner versus outcome syntactic distinction presented above, in conjunction with children’s expectations about agents and actions. These interpretations of rationality and intentionality in particular, especially as they help to distinguish between an agent’s means and their goals, allow for a maximally functional link between manner/means and outcome/goals.

**Rationality as a Source of Meaning**

Introducing the influence of children’s expectations about agents to a mechanism that also considers grammatical structure is a likely model of facilitated word learning. As we have seen, syntax can encode whether the speaker privileges the outcome or manner of an event. An analogue in the realm of infant cognition is grounded in children’s understanding of the distinction between the means an actor uses, and the actor’s ultimate goal. This ability emerges in infants as young as 18-months-old, who selectively perform the *intended* action of an experimenter, both in conditions where the experimenter’s attempts were successful *and* when they were not (Meltzoff, 1995). This not only suggests that infants have the ability to infer the intentions of the adult experimenter, but also that these inferences focus on the goal (i.e. successful outcome) of the action, rather than the means used, in vain or not, to achieve it.

Subsequent work offers further support of infants’ fixation on goals. In one study, infants were habituated to seeing an actor reach toward and grasp one of two objects placed side-by-side on a stage (figure X). During the dishabituation phase, the position of the toys was reversed, and infants saw a series of two subsequent events. In the *path preservation* event, the agent reached for the same location on the stage, thus grasping a new object (given that the objects’ position had been swapped). In contrast, the agent in the *goal preservation* event grasped the same toy as in the habituation phase, but necessarily took a new path to do so. The researchers found that infants as young as 5-months-old looked longer at the event in which the path was preserved (Woodward, 1998). This suggests that even very young infants are viewing agents’ actions in terms of their ultimate goals, and forming related expectations based on this perspective.

These studies demonstrate infants’ expectation that actions are goal-directed, and further, that they are able to interpret this goal even when ineffective means failed to achieve it. However, this does not yet speak to infants’ interpretation of the role of means, specifically, in goal-directed action.

In an attempt to explore the ways in which infants may incorporate expectations about means into a more holistic interpretation of an event, 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; Gergley, 2002). 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 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).

These results were used to propose a naïve theory of rational action, in which children utilize information about the unique constraints of a situation to guide their interpretation of whether or not a particular means was efficient—and thus worthy of imitation, in Gergley’s (2002) study. Similar patterns of rationality judgments have also been observed beyond the scope of imitation, as well. For example, 12-month-olds who see an actor arc their arm over a barrier to reach a ball on the other side are more surprised to see the actor continue to use the novel arm-arc when the barrier is removed than when the actor takes an efficient, direct reach to the ball (Phillips & Wellman, 2005). This also draws the important distinction between infants’ intolerance for novelty and inefficiency. The results from both Gergley (2002) and Phillips and Wellman (2005) suggest that infants are not simply dishabituating to or basing their imitations on novelty, alone. Instead, this novelty is considered within the larger context of what is situationally appropriate.

This leaves open an interesting opportunity to explore exactly what sources of information are used in this situational appraisal. One novel and highly relevant approach was to examine the extent to which language could be used as a source of information as children form expectations about the events they see. This again implicates the ties between language and cognition, specifically with regards to children’s knowledge of syntactic cues and rational goal-directed actions.

**Where Syntax and Rationality Converge**

To examine the relationship between semantic cues conveyed through syntax and infants’ judgments of an actor’s rationality, one study introduced a linguistic context to the classic rational actor imitation paradigm (Kline & Snedeker, 2015). All participants saw the experimenter perform a novel head-touch in the *hands occupied* position, identical to the action featured in Gergley et al (2002). 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”). This structure privileged the outcome of the event (i.e. the toy turning on) as the most essential feature of the action. 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”). Similarly, this structure encoded the manner in which the action was performed as the most essential feature of the event.

The experimenter performed the novel head-touch, 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). If children were indeed using the syntactic frame the novel verb appeared in to inform their rationality judgments, participants in the *manner* condition would use the sentence’s grammatical structure to deduce that the novel head-touch was an essential feature of the event. This would cause them to more closely imitate the experimenter’s novelty, rather than performing the more efficient hand-touch that would be elicited by the situational constraints without the presence of language.

The researchers found that participants in the *outcome* condition performed proportionately more hand-touches when imitating the novel event (Kline & Snedeker, 2015). This result is unsurprising, given that they received both syntactic and observational cues encoding the goal as the essential feature, thus privileging the more efficient means to achieve that goal. 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 source of information called upon to guide children’s interpretations of rationality.

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 shifted 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 to inform their holistic interpretation of an event. 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**

To allow for 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 meant that various patterns could be interpreted more holistically. For example, if participants in the *hands exposed* conditions performed more head-touches than those in the *hands occupied* conditions, this would support a successful replication of Gergley et al.’s (2002) finding that infants systematically use contextual cues to develop naïve theories of rational action. Further, if participants in the *manner* conditions performed more head-touches than those in the *outcome* conditions, this would support two conclusions: first, that the semantic difference between manner versus outcome is salient to infants in our target age range; and second, that this difference is a cue that guides infants’ interpretation of events as they see them occur in the world.

Taken together, these predictions give rise to a sort of cue gradient, in which infants in the *hands 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 *hands 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.

As was the case for 2-year-olds in Kline and Snedeker’s (2015) study, participants’ first contact with the toy was recorded as a measurement of whether they sought to imitate the manner (head-touch) or outcome (hand-touch) of the demonstrated event. While 2-year-olds’ firsts responses proved to be significantly different between conditions, a majority of rational action studies conducted with infants 18-months-old or younger only found a significant difference in response patterns between conditions when participants were allowed to freely explore the toy, performing multiple action attempts (e.g. Kiraly, Csibra & Gergely, 2013). However, it could be the case that this difference between 2-year-olds and younger infants is attributable simply to the size, rather than the nature, of the effect. To explore this possibility more systematically, we introduced an exploration period characterized by the failure of the toy to activate upon first contact.

We predicted that, if participants used linguistic and observational cues to form expectations that privilege *how* the novel event is performed, it is of no consequence if their imitation does not result in the same outcome as when it was demonstrated by the experimenter. In short, so long as they have performed the novel head-touch, they have successfully imitated the novel event. We might then expect to see less interaction with or persistence in trying to activate the toy during exploration, as they already believed they had successfully performed the intended action (i.e. the head-touch). In contrast, if the language and hand cues cause participants to form expectations about the *result* of the novel event—namely, that the toy will light up when touched—then failure of the toy to activate may lead infants to explore the toy for longer, to utilize differing strategies (i.e. manners) to bring about the desired result, or perhaps to express more frustration relative to participants in the manner condition as a direct result of their perceived failure to imitate the novel event. Importantly, this prediction holds regardless of the means used during the first contact with the toy. In order to preserve this, the novel helicopter toy used in Kline and Snedeker (2015) needed to be refabricated.

With the original toy, a participant may roughly “succeed” at activating the helicopter by manually spinning its blades during a hand-touch. While this has the potential to positively reflect our predictions regarding the desires of infants in the outcome-focused conditions to imitate the most effective means to bring about the goal, it prevents our ability to more thoroughly evaluate these expectations by introducing a surprising context in which this goal is not achieved through these efficient means. For this reason, the helicopter was replaced with a clear plastic globe filled with spinning lights that could only be activated by a hidden button concealed on its handle. A large silver dome was also added, positioned opposite the globe on the top of the box surface. This feature was meant to focus infants’ imitation response to one local area, given their familiarity with acting upon buttons. Using this updated toy and the more robust experimental design, Experiment 1 was an attempt to test the extent to which language could shift 18-month-olds’ interpretations of rationality when observing a novel event.

**Method**

**Participants**

Participants were twenty 18-month-olds (range 17 months 3 days to 19 months 3 days; 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 infants 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. A globe was situated a few inches away from a silver 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 infant 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 infant 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 positioned 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 infant 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* and *hands occupied* conditions, the experimenter exclaimed that she was cold, and proceeded to wrap herself up in a blanket made of blue fleece. In the *hands 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 *hands exposed* 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 participant.

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 head touch event, which consisted of the experimenter leaning forward to touch the silver button on top of the toy with her head (*see* Gergley et al., 2002 *for review*) 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, each participant heard the critical sentence a total of four times. The sentence was introduced one final time when the toy was placed within the infants’ reach, and the experimenter prompted the child to *dax (to)* the toy. When the participants 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 infants’ 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. Participants 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 participant’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 participant 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 18-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. Both anecdotal evidence and the findings from previous research (e.g. Kiraly, Csibra & Gergely, 2013; Kline & Snedeker, 2015) suggested that manifestation of infants’ imitation attempts at this age are extremely varied, and 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**

The frequency of each first response type, either head-touch or hand-touch, is presented in Figure X. Contrary to our predictions, all but one infant performed the hand-touch baseline response, regardless of condition. Consequently, participants’ behaviors during the exploration period were not coded as part of this analysis.

Exploratory descriptive analyses of our sample revealed that there were no significant differences in the average age of girls and boys (*t* = 0.28, *p* = 0.78) or between the conditions (*t* = -1.64, *p* = 0.13). Further analyses also suggested that participants’ vocabulary scores were consistent between conditions (*F* = 0.26, *p* = 0.76) and also between genders (*t* = 0.50, *p* = 0.62). However, there was a correlation between participant age and vocabulary score, *r*(19) = 0.46, *p* = 0.05.

**Discussion**

Experiment 1 was designed to test the relationship between the structures in which a novel verb is presented, and infants’ rationality judgments about the novel event the verb was meant to describe. If infants used both language and action cues to inform their interpretation of the event, we predicted that participants across conditions would form different judgments about the experimenter’s rationality. Infants in the *outcome + hands occupied* condition would use these language and action cues to contextually justify the demonstrator’s novel means, but would nevertheless opt to use the more efficient hand-touch when imitating the event themselves. However, infants in the *manner + hands exposed* condition would use these cues to infer that the demonstrator’s novel means were essential to the event itself, and thus seek to imitate that novelty, as well.

Instead, we found that virtually all participants performed a hand-touch imitation. Given the homogeneity of response type, the exploration portion of the experiment was not coded for inclusion in the final analyses. Without producing virtually any head-touch responses, we cannot be certain whether patterns found using this paradigm are a true representation of infants’ cognitive and linguistic abilities, or simply the result of an uninformative methodological flaw. In either case, it is unlikely that this unexpected finding can be attributed to any sampling errors or characteristics of our participants. Exploratory descriptive analyses showed there to be no relationship between participants’ age, gender or vocabulary size and the condition they were placed into. The only observed correlation—between age and vocabulary size—is neither surprising nor consequential. We would surely expect older infants to have larger vocabularies, and this effect is neutralized by the fact that each condition had a similar distribution of ages, and therefore also of vocabulary sizes.

Amongst the possibilities that remained was the claim that our paradigm was, in fact, sensitive to the effects we set out to measure, and 18-month-olds’ desire to use the more efficient means is an accurate reflection of their prioritization of goals. This would, however stand in opposition to a large body of literature that characterizes the sensitivity of infants’ event perception to rich sources of contextual information. Our findings were therefore more likely to be the result of some form of procedural or material limitation. Accounting for these shortcomings was the focus of Experiment 2.

**Experiment 2**

The same 2 x 2 (hands x language) design used in the first experiment was also used for Experiment 2. However, the materials and procedure of the second experiment were modified with the goal of eliciting head-touches in addition to the previously recorded hand-touches. If our methods could be changed such that we are able to elicit both response types, this would suggest that, within the context of our paradigm, there is some circumstance in which infants are willing to perform a goal-directed action using novel means. Only then would we be able to draw conclusions as to what exactly about these circumstances cause infants to deem the otherwise indirect action as being necessary.

To that end, the first consideration made was in regards to potential physical constraints within the paradigm. Perhaps it is the case that 18-month-olds 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, has featured infants as young as 14-months-old successfully completing head-touches (see also Paulus, Hunnius, Vissers & Bekkering, 2011). This would suggest that our participants have the motor skills necessary to perform a head-touch, and thus the novel lean was kept as our indirect manner.

The next consideration made was with regards to the toy, itself. After thorough review of previous rational actor imitation paradigm studies that used any remotely similar type of light box stimulus, two key alterations became evidently necessary. First was to change the actual dimensions of the toy. Most toys used in the other studies were mounted on boxes ranging in height from a mere 4.5 to 6cm, 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 demonstrating 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). However, evidence from Experiment 1 suggests that this causal link may have been too opaque for 18-month-olds, given that over half (55%) of participants in the first experiment made contact with the globe, rather than the button, as their first imitation response (*figure X*). 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.

In addition to changes made to the toy, a longer, more effective warm-up period was implemented in Experiment 2. The added emphasis on ensuring all participants were sufficiently comfortable before the start of the critical trial served two functions. The first was to minimize the number of tested infants who became too fussy or uninterested in completing the trial, and would consequently be ineligible for inclusion in the final sample and analyses. The second function was more directly related to the goal and hypotheses of Experiment 2. A participant’s comfort and willingness to engage may have a direct effect on their response type, insofar as performing a head-touch requires the infants to physically separate from their parent. If a child is feeling somewhat unsettled in the situation, they may seek to remain in physical contact with their parent, and opt to perform a hand-touch, which they can complete while still securely attached to mom or dad. Thus, if we want to ensure that our paradigm is capable of eliciting head-touches, we needed to create an environment in which participants were both physically able and emotionally willing to perform the novel action.

Participants’ first responses were coded according to the criteria set forth in Experiment 1. An additional coding scheme was devised for use in Experiment 2 to more systematically parse the full range of participants’ exploratory behaviors. This coding scheme was organized hierarchically, based on the level of detail encoded within the data point or interval. At the two coarsest levels, videos were coded for the durations of time infants were engaging with the toy, and intentionally making physical contact with the toy. While these measures captured broad approximations of infants’ behaviors, they were crucial in determining the validity of our prediction that infants who had formed expectations specifically regarding the *outcome* of the event would be more persistent in turning on the toy when it failed to operate (i.e. engage more with the toy and/or perform more intentional body actions).

The next level of measurement captured individual instances of hand- or head-touches. These annotations were meant to be a detailed behavioral record for each participant that specified the number and kind of actions performed. This would allow us to measure the degree to which a participant’s condition is able to account for the variations in response patterns. The duration of each segment of the trial—warm-up, demonstration, first response, exploration and total trial length—was recorded as well, for the purposes of exploratory descriptive analyses.

Using this updated coding schema and toy, Experiment 2 was an attempt to (1) confirm that head-touches could be elicited using this particular paradigm; and (2) test whether hand condition, language condition or some interaction between the two could systematically account for the variance in participants’ imitation response patterns.

**Method**

**Participants**

Participants were forty-five 18-month-olds (range 17 months 4 days to 18 months 29 days; 20 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 infants 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 infants’ 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 exposed* and *hands occupied* conditions.

The lights in the testing room were slightly dimmed to increase the salience of the toy’s activation, and a small camcorder was positioned facing directly perpendicular to the participant to record their interaction with the toy.

**Procedure**

Families were greeted upon arrival to the lab. As in Experiment 1, the experimenter engaged the infant in interactive free play in the lobby while giving parents instructions on how to neutrally respond to their child during the exploration period of the study. Given that a lack of interest was a meaningful measure of engagement, parents were also reminded not to specifically direct their child’s attention to the toy, and to only reposition the toy in front of their child if the infant was making a clear indication of their desire to reach the toy in the event that it had accidentally fallen off or been pushed to the far side of the table. Once all forms were completed, the families were escorted to the testing room for an extended warm-up period.

Unlike in Experiment 1, a number of toys were displayed on the table at the start of the experiment. These included several small stuffed animals, a set of colored building blocks, and a textured red ball. The experimenter engaged in free play with the infant 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 infant 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 infant. The ball was put away once the experimenter felt that the participant was sufficiently familiar with both her and the puppet, at which point the Simon-Says-like warm-up task from Experiment 1 was initiated.

In contrast to the first experiment, parents in Experiment 2 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!” The puppet was put away at the end of the warm-up trials before the novel toy was introduced and demonstrated by the experimenter exactly as it had been in Experiment 1.

Further following the original procedure, the exploration period began after the infant made first contact with the toy, to which the experimenter responded with the neutral, yet enthusiastic reply, “Okay! Now you can play.” After the duration of the exploration period, the experimenter returned to the table and activated the toy, enthusiastically praising the infant. The session was then ended, and families were debriefed and thanked for participating.

**Coding**

All sessions were videotaped in order to accurately assess each participant’s interactions with the toy. The experimenter coded first contact with the toy live during each session. Using the same criteria as in Experiment 1, first contact was labeled as either a “hand-touch,” “head-touch,” “N/A” (for no response), or “fuss out,” when a participant was unable to complete the experiment. This and all other behavioral measures were catalogued using the video annotation tool VCode, which allowed us to track both single-point and duration events during the trials (*figure X*).

The duration of each major segment of the trial as recorded in the first round of VCode annotation. These segments included the total trial length, warm-up period duration, novel action demonstration duration, the first response window, and the length of the exploration period. The first response window and exploration period were coded a second time, now indicating both the onset and type of the first response, as well as all other individual instances of head- or hand-touches. After reviewing a sampling of the trial videos, a half-second time delay appeared to be a consistent quantification of the separation between hand-touch attempts. Accordingly, for any hand contact made during the exploration period to be coded as a proper hand-touch, there must have been a half-second delay between its onset and the offset of any hand contact that preceded it.

In the final phase of coding, the durations of an additional set of measures were recorded. First was each participant’s *engagement*, which characterized any combination of (1) looking at the toy; (2) touching the toy; or (3) discussing the toy in some capacity (e.g. talking about the toy directly, asking for a parent’s help to fix it, etc.). The second measure was a slightly more specific record of *engagement* during the exploration period. These annotations, labeled *intentional body actions* (IBAs), captured intervals in which the participant was (1) making physical contact with the toy; and (2) looking at the toy *at the onset* (but not necessarily for all) of the contact period.

**Results**

Descriptive analyses revealed there to be no difference between conditions in average participant age (*df* = 3, *F* = 0.08, *p* = 0.97) or vocabulary size (*df* = 3, *F* = 0.78, *p* = 0.51). An analysis of variance (ANOVA) was used to compare average duration length for each segment of the experiment. There were no significant differences in total trial length (*df* = 3, *F* = 0.85, *p* = 0.47), warm-up period duration (*df* = 3, *F* = 0.66, *p* = 0.58), length of the first response window (*df* = 3, *F* = 1.19, *p* = 0.33), or exploration period duration (*df* = 3, *F* = 1.00, *p* = 0.40) between conditions. However, there was a significant difference in the average duration of the demonstration period, *df* = 3, *F* = 3.26, *p* = 0.03. Further analyses revealed that participants in the *hands exposed + manner* condition witnessed the longest demonstration of the novel event, lasting on average 48.5 seconds, compared to an average of 43.7 seconds across the other three conditions.

Participants’ first imitation responses—either head-touch or hand-touch—are presented in figure X. A logistic regression revealed there to be no change in response type as a function of either hand condition (*b* = -0.94, *p* = 0.20) or language condition (*b* = -1.10, *p* = 0.13), net of the other. We then analyzed the response behaviors of participants during the exploration period (figure X). The number of head-touches was not significantly predicted by hand condition (*b* = -0.04, *p* = 0.95), language condition (*b* = 0.21, *p* = 0.74), or any interaction between the two (*b* = 0.31, *p* = 0.80).

Further decomposing the interaction, (SIMPLE SLOPES) results showed X relationship between hand condition and the number of head-touches performed during the exploration period at the *outcome* (“dax”) level of language [stats], and X relationship between hand condition and the head-touches performed during exploration at the *manner* (“dax to”) level of language condition [stats]. A subsequent analysis of variance (ANOVA) revealed that the total number of actions performed (i.e. inclusive of both head-touches and hand-touches) was not significantly different between conditions, *df* = 3, *F* = 0.62, *p* = 0.61.

Next we tested the prediction that forming expectations about the outcome of the novel event, specifically, would lead participants to explore the toy more persistently when this expectation was violated. The amount of time spent engaging with the toy during the exploration period was not significantly predicted by condition, *b* = -3.85, *p* = 0.73. However, the ability of each condition to predict the amount of time spent physically interacting with the toy during the exploration period (measured as IBAs) was trending toward significance, *b* = 10.41, *p* = 0.09, though in the opposite direction as predicted.

**Discussion**

1. Reminder: what we expected, what we found
2. Coding scheme: strengths & weaknesses
   1. The original idea: hierarchy to capture multiple levels of interactions
   2. BUT no trends either in line with OR strongly against our hyp
3. Paradigm: strengths & weaknesses
   1. We got head touches, yay!
   2. Remaining possible limitations
      1. Maybe hands occ isn’t convincing enough / not natural
         1. Suggestions
      2. Other?

**General Discussion**

*Now looking at both sets of results*

* 1. Theory: strengths & weaknesses
     1. Contextualizing our results: reminder of why we had reason to believe the effect might exist
     2. Gergley’s failure, too...
        1. WHY: 18mos & event cognition
        2. WHY: 18mos & language

1. Other smart things