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, Laura, Sarah

LDS Grant Info

Mellon Funding

Mom

**Table of Contents**

**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**

**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. To that end, 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. Gergely, Bekkering & Király, 2002). At the surface, there is an intuitive benefit in concurrently learning about actions and the structure of the verbs meant to describe them.

If the two are connected, one possible conclusion could be that each domain draws information from the other to facilitate their own, respective processes. This would suggest a certain level of flexibility within each domain, insofar as this would require them to incorporate external information into their own, internal schemas. Implicitly, this claim also suggests that there is some inherent benefit to integrating aspects of the two domains. Perhaps doing so decreases the cognitive cost of computing logical inferences (e.g. reasoning about why an agent performed an action or described an event in a particular way). Their connection could also suggest that, on their own, each individual domain lacks the sufficient information or processing ability to tackle the cognitive task at hand (e.g. perhaps we need language as a medium to reason about events, or we need to form expectations about the events to reason about what an agent is referring to). Alternate interpretations could also posit that there exists a single source of information from which both domains are drawing knowledge (e.g. the domain-general concept of “agent”), or that one domain contains a specific concept that is being recruited by the second domain (e.g. the event-related concept of “agent” used in a linguistic context; see figure X).

Again, there is also the possibility that representations within language and event cognition are wholly distinct, simply maturing in parallel during development. If the encapsulated, domain-specific account were to be true, it would require that each domain has a unique representation of “agent” that allows the child to infer (1) how agents speak about themselves and others (in the domain of language) and (2) what agents tend to do and why (in the domain of rational events).

Each possible account has its theoretical benefits and drawbacks. On the one hand, having a generalized concept of “agent” would allow for computational simplicity (i.e. only a single concept is needed), but may render this concept underspecified within each of the domains it is used. On the other hand, having unique, domain-specific concepts of “agent” could be computationally redundant (i.e. storing replicas of near-identical concepts), but this may be overshadowed by the power within the nuances of each concept’s domain specificity. But parsing these implications necessarily requires first observing the ways in which the domains interact. How exactly could another domain of knowledge facilitate children’s language acquisition?

One argument 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. Yet, children’s ability to develop adult-like grammars in spite of this could be used to motivate the suggestion that they have an innate language learning mechanism, or that they make use of cognitive constraints that facilitate language acquisition (e.g. Jackendoff, 1975).

Such constraints may also help children overcome the most fundamental challenge in language acquisition: mapping words to their referents. This process 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”), subsets (e.g. “animal,” “dog,” “Fido”), or verbs that extend across agents and manners (e.g. “Jane runs slowly” vs. “the cheetah runs quickly”).

If it were the case that children relied only on statistical observations of words said in the presence of their referents, there would be large swaths of the adult lexicon that remain unlearned. There will never be a case where a child observes a dog that is not an animal or an apple that is not a fruit. Children who hear imperatives would be baffled. How could they understand adults who are referring to the exact thing that is *not* present in the conceptual space (e.g. “Brush your teeth!” when no teeth-brushing is occurring)? Simple associations could also not account for how we come to discuss *features* of objects or events. The “associationist” child would forever think that “fur” referred to the whole creature—rather than just its coat—or that “splash” described the entire event of jumping off an elevated surface and coming to rest in a body of water—rather than just the displacement of that water.

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 further 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 constraints 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. What we have seen thus far, then, is a strong case for the theoretical benefit of a link between language and other domains of knowledge that privilege certain concepts new words and phrases can map to. But our original question remains unanswered: how exactly does this manifest within the domains of verb structure and event knowledge, if at all?

To explore this question further, we will begin by examining how children may derive semantic content from syntax, and address which necessary components of verb meaning are not sufficiently conveyed through language, alone. Bridging this gap between what is conveyed in a verb’s structure and what is understood about an event will lead us to the developmental literature on children’s knowledge about agents. Here, we will discuss what a child needs to know about agents acting on object in order for them to successfully use their knowledge of events to scaffold their understanding of the verbs that describe them. The manifestations of this dependency between syntactic cues and behavioral cues will then be borne out in a series of rational actor imitation paradigms. These paradigms allow us to test how children integrate their expectations about what an agent is saying, and what they are doing. Ultimately, we will use the findings of these experiments as a lens to make sense of the mechanisms and concepts that facilitate this cross-domain reasoning.

**Syntax as a Source of Meaning**

As we have seen, children rely on a number of cues to help infer the meaning of a novel word. These cues can be non-linguistic information, such as joint-attention (NAME,XXXX), but may also be purely linguistic in form. For example, 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 will look reliably longer at a video depicting a causal event after hearing a transitive sentence (Naigles, 1990). This remains true even when all of the noun phrases are pronouns, making them structurally more ambiguous (Fisher, 1996), or 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 19-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).

This evidence suggests that children are able to learn about these particular features, but what are the actual constituents of meaning that are conveyed through these structures? For example, transitive structures that encode causality help the child deduce who is doing the causing, what that cause is, and who experiences the effect. In the sentence, “John pushed Mark,” we are able to infer who is the “causer” (*John*), who is the “effectee” (*Mark*) and what the relationship is between them (*pushing*). Another proposed structure conveys important information about *how* something is done (manner) or the *result* of the thing done (outcome). 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 (i.e. the table is 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 (Levin & Rappaport Hovav, XXXX). 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, there are equally supported arguments as to why syntax alone is not a sufficient source of semantic information.

One 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 be sure, syntactic structure conveys important information about the type of verb it is likely to contain. But how might a child bridge the gap between inferring what a novel verb is *likely* to be, and knowing what it actually *is?* One source of secondary information may simply be observation, though in a sense more complex than proposed by association theories or just “looking at what lies on the other side of the camera” in our zoom lens metaphor. In this context, observation may be thought of as yet another constraint on possible verb meaning that originates from non-linguistic knowledge about the world.

For example, when we talk about actions and events (i.e. use verbs), it may be useful to have expectations about what an agent is, what sorts of things they are likely to do, and why. All of this information, when combined with verb syntax, may bring the child closer to understanding the full semantic contents of a verb. To further illustrate: if a child is presented with a novel verb in a structure that encodes means or results of an event (as in the case of the manner vs. outcome distinction), it is necessary for that child to be able to distinguish between these two components of an event. While this distinction may appear to be intuitive or trivial at best, developmental literature has studied at length when and how children are able to reason about events in this way. Thus, we turn next to exploring the content of these expectations, and how they could be used to facilitate understanding the meaning of a novel verb.

**Rationality as a Source of Meaning**

In order for knowledge about events to be a useful source of information when inferring the meaning of a manner or outcome verb, there must be some form of conceptual mapping between the linguistic and behavioral domains. 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. Take for example the following scenario: say Sally bumps into a table that has a lamp on it, causing the lamp to fall over and break. We could describe this event in a number of ways: “Sally hit the table (causing the lamp to break)”, “Sally broke the lamp (by hitting the table)”, and so on. An adult speaker recognizes that each of these sentences are true, and refer to different perspectives on the same event: the means and outcome, respectively. It could be the case that children represent this event as one conceptual unit, roughly: [events where things are *HIT* + *BROKEN*].

If this were the case, conceptual knowledge in the domain of event cognition would be functionally useless in helping children distinguish between these two types of sentence structures. However, evidence suggests that infants as young as 18-months-old have the ability to represent an agent’s manner as a distinct concept, separate from the agent’s goal. In one study, infants 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 suggests that the infants are not representing the entire event as one concept (e.g. [EVENT]*successful* or [EVENT]*unsuccessful*), but rather as its component parts (e.g. [MEANS]*successful*, [MEANS]*unsuccessful*, [GOAL]). Moreover, this supports that infants have the ability to infer the intentions of the adult experimenter, and 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 (Woodward, 1998; *figure* 1). 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). If infants viewing these displays simply attended to novelty, rather than to more abstract expectations about goal-directedness, we would have expected participants to look equally long when the toy was in a new place as when a new path was taken. The fact that they were selectively surprised that an agent would have a new goal while indifferent about that agent’s means suggests that even very young infants are viewing agents’ actions in terms of their ultimate goals, and forming related expectations based on this perspective.

While these studies offer a clearer picture of how infants represent goals, they do not speak to infants’ interpretation of the role of *means* in goal-directed action. Based on the evidence presented, it could be the case that means have little cognitive significance: that they quite literally are simply a way to achieve a goal. However, subsequent research provides evidence that the means of an event are functional representations in their own right that contribute to the broader interpretation of an event.

In one famous example, 14-month-olds saw an unfamiliar event in which an experimenter achieved a goal (turning on a light box) by utilizing a novel means (leaning forward to touch her forehead against a light box; Gergely *et al.*, 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 Gergely 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 (Gergely *et al*., 2002).

These results were used by Gergely to propose that infants have a naïve theory of rational action, in which they 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 Gergely *et al.*’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 again draws the same important distinction between infants’ intolerance for novelty and inefficiency we saw in Woodward’s (1998) study: the results from both Gergely (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. However, this leaves open an interesting opportunity uncover exactly what sources of information are used in this situational appraisal. One approach to explore this would be to directly examine the extent to which language can be used as a source of information as children form expectations about the events they see.

**Where Syntax and Rationality Converge**

We know that even very young infants display robust, context-sensitive expectations about agents and their goals. Moreover, we have seen evidence that infants are also able to use a verb’s syntactic frame to constrain its meaning. However, possible mechanisms that directly integrate these domains are limited. Chen and Waxman (2013) demonstrated that use of a novel word can lead infants to more frequently imitate a novel means than when their attention is simply drawn to the event. But this account stops short of testing the effect of syntax directly by introducing the same novel verb in contrasting subcategorization frames. On the other hand, Yuan *et al.* (2012) highlight a contrast between different structures that contain the same novel verb, yet examine a causal distinction (transitivity) that need not require children to use their knowledge about goal-directed events.

What is missing then is a paradigm that explicitly examines a mapping between manner/outcome cues in syntactic frames and children’s expectations about rational action. To address this, one study introduced a linguistic context to the classic rational actor imitation paradigm (Kline & Snedeker, 2015). In this study, all participants saw the experimenter perform a novel head-touch in the *hands occupied* position, identical to the action featured in Gergely *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 language* 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* *language* 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 language* 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 manner, 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* *language* 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* *language* 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 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**

A 2 x 2 (hands x language) between-subjects design was used. This allowed us to explore two sets of predictions. First, if 18-month-olds use situational constraints to inform their judgments of an actor’s rationality, we would expect to see more hand-touch imitations in the *hands occupied* condition, and more head-touch imitations in the *hands exposed* condition. This would suggest that, like younger infants (cf. Gergely *et al*., 2002), 18-month-olds in the *hands occupied* condition attribute the experimenter’s novel action to the fact that her hands are unavailable to her, thus prompting them to use the more efficient hand-touch to achieve the goal. Moreover, this would also suggest that 18-month-olds in the *hands exposed* condition interpret the novel action as being voluntary, and thus a more essential component of the event, resulting in a higher proportion of novel head-touch imitation responses.

The second set of predictions concerns infants’ use of linguistic cues. If 18-month-olds also simultaneously use syntactic information to deduce whether the novel verb refers to either the means or the outcome of the event, we would expect to see participants in the *outcome language* condition to seek the most efficient means to imitate the outcome, thus performing more hand-touches. In contrast, we would expect participants in the *manner language* condition to perform more head-touch responses as a means to more faithfully imitate the novel manner demonstrated by the experimenter.

Taken together, these predictions give rise to a sort of cue gradient, in which infants in the *hands exposed + manner* *language* 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* *language* 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. However, if we observe no variance in response type across the conditions, this would suggest that our particular paradigm is either unsuccessful in properly conveying these linguistic and behavioral cues, or unable to accurately measure participants’ sensitivity to them. In this case, answering our experimental question would require necessary improvements to our methodology.

We also sought to explore the extent to which infants’ judgments informed robust expectations about the novel event. To test this, we introduced an exploration period characterized by the failure of the toy to activate upon first contact. If participants in the *manner language* and *hands exposed* conditions used linguistic and observational cues to form expectations that privilege *how* the novel event is performed, we would expect to see less interaction with or persistence in trying to activate the toy during exploration, as they may already believe they have successfully performed the intended action (i.e. the head-touch). Essentially, it is of no consequence if their imitation does not result in the same outcome as when the experimenter demonstrated the action: so long as they have performed the novel head-touch, they have successfully imitated the novel event.

In contrast, if the language and hand cues cause participants in the *outcome language* and *hands occupied* conditions 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 other conditions. 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. If they are using both behavioral and linguistic cues to inform their interpretations, we expect participants in the *hands exposed* and *manner language* conditions to perform the most novel head-touch imitations. Further, we expect participants in the *hands occupied* and *outcome language* conditions to privilege the more efficient means to achieve the goal, thus performing more hand-touch imitation responses. Extending the previous work of Kline and Snedeker (2015) to younger infants allows us to explore the potential developmental trajectory of the ability to integrate language and behavioral cues. Our particular methodology also allows us to test a number of additional predictions, such as the effect of conflicting cues or a violation of expectation, which leads to a more comprehensive characterization of this mechanism for cross-domain reasoning.

**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 2*). The globe’s handle was concealed within the box, and was wired to a button to facilitate hands-free operation of the spinning lights.

**Procedure**

Families were greeted upon arrival to the lab, where the experimenter engaged the child in interactive free play in the lobby. During this time, the experimenter reviewed the procedure with each family, and provided instructions to parents on how to interact with their child during the experiment. To reduce any potential for biases, parents were 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.

Once 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. A small camcorder was positioned facing directly perpendicular to the table to record the infant’s interactions with the toy.

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* *language* condition or, “Look! I’m going to *dax* my toy!” in the *outcome language* 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* Gergely 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* 3).

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 participants “they could play” or that “it was their turn” before walking off to another corner of the room and shuffling papers to look preoccupied. 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. Király, 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.

We intended to generate an exploration period coding scheme based on a subset of participants’ videotaped sessions. However, we were unable to develop an index of hypothesis-relevant measures due to a lack of variability in participants’ response types (see Results below for more details).

**Results**

The frequency of each first response type, either head-touch or hand-touch, is presented in Figure 4. Contrary to our predictions, all but one infant performed the hand-touch baseline response. A chi-squared test with expected values of 0 confirmed that our results were not significantly different from if we had observed no head-touches at all, *X2*(1, *N* = 20) = 0.05, *p* = 0.83. 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 across conditions (*F* = 0.26, *p* = 0.76) and also across 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 in the *outcome language + hands occupied* condition perform more hand-touches when imitating the event, we predicted that they would do so because they were using these language and action cues to contextually justify the demonstrator’s novel means. Given that they were not under the same situational constraints, they could satisfy their expectations of efficiency by utilizing the more efficient hand-touch imitation response. In contrast, if infants in the *manner language + hands exposed* condition perform more novel head-touches, this would suggest that they were using these cues to infer that the demonstrator’s novel means were essential to the event itself.

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.

To that end, the success of our paradigm hinged on three major components: (1) the ability of our methods to convey the necessary cues; (2) the ability of our methods to measure infants’ sensitivity to these cues; and (3) the willingness of our participants to engage with and imitate the experimenter. Perhaps the size, shape or arrangement of the toy within our paradigm limited the information conveyed to the participants, or unrepresentatively constrained their ability to interact with they toy at test. It could also have been the case that the toy, or even the experimenter, was not sufficiently engaging to the infants. Given that our behavioral metric was imitation, our findings could simply suggest that participants had no social desire to imitate the experimenter, rather than speak to their ability to interpret behavioral and linguistic cues.

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 observed correlation between age and vocabulary size likely did not influence our findings, as ages—and thus vocabulary sizes—were normally distributed across conditions (figure 5). Instead, this serves as a successful replication of Fenson and his colleagues (2000) within a markedly smaller age range of only two months.

Beyond the methodological and sampling influences previously discussed, an alternate approach to interpreting our data would be to suggest that our paradigm was valid, and our findings are therefore an accurate reflection of 18-month-olds’ cognition. However, the lack of variability in imitation response type strongly suggests that this is not the case. In the subsequent experiment, our focus thus shifted to eliciting novel head-touches in addition to hand-touches.

To do this, we accounted for the major methodological limitations present in Experiment 1: the toy was lowered to more infant-friendly height, and also to increase the salience of the head-touch action; we reconstructed the toy to appear more simple to activate; and we modified the warm-up period of each trial to promote infants’ engagement and social comfort with the experimenter. If these changes lead to a number of infants successfully performing head-touches, this would confirm that our paradigm is conveying the intended linguistic and behavioral information. Moreover, this would also suggest that our paradigm is able to measure participants’ sensitivity to these cues. We would then be able to test whether 18-month-olds’ expectations about rational action can be shifted by language, allowing us to draw inferences about how children come to utilize this information across domains of knowledge.

**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 Gergely *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 6*). 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 7*). 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!” At the end of warm-up period the puppet was put away, and the procedure continued identically to Experiment 1.

The experimenter introduced the novel toy before exclaiming that she was cold and wrapping herself up in a blanket. In the *hands exposed* conditions, she draped the blanket over her shoulders, with her hands placed on either side of the toy still visible to the infant. In the *hands occupied* conditions, the experimenter had her hands concealed underneath the blanket, out of sight of the participant. She then introduced the critical sentence: “I’m gonna *dax* my toy” in the *outcome language* condition, and “I’m gonna *dax to* my toy” in the *manner language* condition.

The experimenter then performed the novel head touch, activating the toy when her forehead reached the globe, which created the illusion that physical contact caused the lights to illuminate and spin. She then repeated the critical sentence to describe the event, before performing the action/sentence sequence a second time. By the end of the demonstration, each participant saw the novel event twice and heard the critical sentence four times. The infants heard the critical sentence one final time when the experimenter prompted them to *dax*(*to*) the toy, eliciting their first imitation response.

**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 (Hagedorn, Hailpern, & Karahalios, 2008; *figure 8*). The specific measures coded during the exploration period of Experiment 2 can be found in Table 1.

**Results**

Participants’ first imitation responses—either head-touch or hand-touch—are presented in *figure* 9. A logistic regression revealed there were no changes 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* 10). The main effect of hand condition did not significantly predict the number of head-touches performed during the exploration period, net of language condition (*b* = -0.04, *p* = 0.95). Further, the main effect of language condition was also not a significant predictor of head-touches, net of hand condition (*b* = 0.21, *p* = 0.74). There was also no interaction between hand condition and language condition, net of the main effects (*b* = 0.31, *p* = 0.80). A linear regression also revealed that neither the main effects of hand or language condition were a reliable predictor of the total number of actions performed during the exploration period, *b =* 0.40, *p* = 0.73 and *b* = -0.14, *p* = 0.91, respectively..

Our next set of analyses concerned whether infants in different conditions formed different expectations about the outcome of the novel event, which we measured by recording the total time spent engaging with the toy during the exploration period. If infants formed strong expectations about the outcome, they may explore the toy more persistently (i.e. longer) when this expectation was violated. A simple linear regression showed that 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 interaction between hand and language condition was marginally able to predict the amount of time physically interacting with the toy during the exploration period, measured as IBAs, net of the main effects, *b =* 10.41, *p* = 0.09 (*figure* 11). Neither the main effect of hands or language was significant net of each other (*b =* 4.39, *p =* 0.16; *b =* 4.88, *p =* 0.12, respectively).

Exploratory analyses showed there were no differences across condition in average participant age (*F*(3,40) = 0.08, *p* = 0.97) or vocabulary size (*F*(3,40) = 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 (*F*(3,40) = 0.85, *p* = 0.47), warm-up period duration (*F*(3,40) = 0.66, *p* = 0.58), length of the first response window (*F*(3,40) = 1.19, *p* = 0.33), or exploration period duration (*F*(3,40) = 1.00, *p* = 0.40) across conditions. However, there was a significant difference in the average duration of the demonstration period, *F*(3,40) = 3.26, *p* = 0.03. An analysis of variance (ANOVA) revealed that participants in the *hands exposed + manner language* condition witnessed the longest demonstration of the novel event (*figure 12*). The average length of the demonstration in this condition was 48.5 seconds. An independent samples T-test confirmed this to be significantly longer than the average demonstration across the other three conditions (43.7s), *t* = 3.02, *p* < 0.05.

**Discussion**

Experiment 2 was designed to explore the potential ability of language to influence infants’ interpretations of a goal-directed event. If infants rely on linguistic cues in addition to perceived situational constraints when forming their judgments of an actor’s rationality, we hypothesized that infants who received both manner-encoding syntax cues and means-focused behavioral cues would be more likely to imitate the novel head-touch performed by the experimenter. In contrast, their peers who received outcome-encoding syntax cues and goals-focused behavioral cues would opt to perform the more efficient hand-touch as a means to imitate the outcome of the experimenter’s goal-directed action. Contrary to our predictions, infants’ condition—and thus, the language and behavioral cues they received—did not reliably predict their mode of imitation.

To fully explore the manifestation of this proposed relationship between syntactic cues and rational action, we also hypothesized that infants in the *hands occupied + outcome language* condition would engage with the toy for a larger proportion of the exploration period following the failed activation of the toy. This prediction arose from an extension of the original hypothesis, which suggested that these infants, but not infants in the *hands exposed + manner language* condition, formulate specific expectations about the outcome of the novel event. When their imitation response fails to activate the toy, this violates their expectations about the outcome, and thus they may be more persistent in attempting to achieve their goal (i.e. turning on the toy). Again, infants’ condition failed to account for the time spent engaging with the toy during the exploration period. Participant condition was also unable to predict our other, more detailed measures, including amount of time spent attempting to operate the toy via physical means (i.e. *intentional body actions*) and the specific number of head-touches, hand-touches, or total actions performed during exploration.

Unlike in Experiment 1, these findings are not likely to be the result of any limitations of the toy, itself. By lowering the toy’s dimensions and introducing a simpler causal model for activation, we were able to successfully elicit head-touches in addition to hand-touches. This was crucial to confirming our particular toy as a valid tool to measure infants’ imitation patterns. Though despite the variability in our data, these patterns could not be explained by infants’ exposure to situational constraints (cf. Gergely *et al.,* 2002), language cues (cf. Chen & Waxman, 2012), or any interaction between the two (cf. Kline & Snedeker, 2015).

This raises interesting questions regarding outstanding limitations of our particular study design. Given the normal, randomized distribution of our participants across conditions, a possible explanation of the data would necessarily need to extend beyond any sampling coincidences. Two possible critiques concern (1) how we *present* the relevant cues and (2) how we *measure* participants’ responses. In regards to measurement, a number of considerations must be made when developing an appropriate coding scheme. As was previously mentioned, the measures of interest were defined using criteria that captured broad behavioral patterns as well as detailed indices of individual actions. The intent was to characterize any possible influence in whatever level of specificity it was evident. Further, our coding criteria closely modeled those used in previous rational actor imitation paradigms (see Appendix X). If the previous findings were stable, using the same measures should be sufficient to at the very least replicate pattern differences between the *hands exposed* versus *hands occupied* conditions. However, this failure suggests that our null result is likely the product of something beyond inaccurate or improperly calibrated criteria.

Thus, this still leaves open the question of whether our presentation of the situational cues served their intended function. Previous research has shown that 12-month-old infants selectively imitate the novel head-touch of an experimenter only when her hands are *restrained* (by being tied to her chair), not simply when they were “voluntarily” occupied (as in the case of holding a blanket around her shoulders; Zmyj, Daum, & Aschersleben, 2009). This would suggest that infants require a certain degree of plausibility in order to consider an experimenter’s hands to be “occupied” in earnest. In the case of Gergely *et al.*’s (2002) participants, this criterion could have been met when the experimenter exclaimed that she was cold, thus prompting her to wrap herself in the blanket. This emotionally valenced component was not presented to the 12-month-olds, and could represent a sufficient indicator of “necessity” to 14-month-olds. It may follow then that even the physical and emotional cues combined do not convey a cue that is sufficiently believable to our 18-month-old participants.

To rule out the possibility that our sample was comprised of skeptics, future iterations of our study should include a *language-free* condition. Doing this would establish a baseline representation of how18-month-olds interpret the rationality of the experimenter’s actions. If they replicate the patterns found in Gergely *et al.* (2002), this would suggest that the *hands-occupied* condition is a sufficiently plausible constraint. Failing to replicate these findings would necessarily call into question the validity of the rational actor imitation paradigm for use with 18-month-old participants. In this case, future iterations would benefit from utilizing other means of assessing infants’ judgments of rationality, such as preferential looking (e.g. NAME, 0000) or performance in role-playing paradigms (e.g. Schwier, Van Maanen, Carpenter & Tomasello, 2006).

However, exploring these limitations only addresses one subset of the possible sources of our null results—namely, that our findings indicate some sort of methodological shortcoming in our experimental design. Another possibility is that our findings accurately characterize the relationship between 18-month-olds’ representations of syntax and rational action. Depending on the specific mode of interpretation, this could lead to a number of claims, some more or less supported. Perhaps 18-month-olds are not sensitive to the linguistic cues in the same way that adults are, or that they are able to recognize these cues, but they are not relevant within the context of a rational imitation task. A parallel possibility exists for the behavioral cues: it could be that 18-month-olds represent rationality differently, at least in this context.

Maybe their performance is not the result of any uncharacteristic representation of either the syntactic or behavioral information, but rather due to a failed mapping between the two. Or perhaps 18-month-olds are fully capable of integrating these domains of knowledge, but do not do so in this context (e.g. because it is not meaningful; or it does not conjure up the simultaneous representations that comprise this link; or it does not kick-start the mechanism that connects the two domains, etc.). Over and above any particular nuances in possible interpretations, a fundamental challenge awaits anyone who wishes to present these null findings as positive evidence: how do we account for instances of non-linear developmental trajectory?

Linguistic and cognitive development are conceived of as processes by which some trait (e.g. particular knowledge, the functioning of a particular mechanism, etc.) “increases” on some dimension (e.g. accuracy, complexity, robustness of extension, etc.) as a function of some predictor (e.g. age or experience). For example, we have discussed evidence that suggests infants *younger* than 18-months-old are able to use situational constraints to interpret rational action (e.g. Gergely *et al.*, 2002) and are also able to use syntactic cues to bootstrap meaning (e.g. Yuan *et al*., 2012). Infants *older* than 18-months develop—in the colloquial sense of the word—the ability to integrate these cues as a means to inform higher-level interpretations of novel events and the way adults speak about them (e.g. Kline & Snedeker, 2015). We might then expect that infants in this middle stage, our 18-month-olds, would express abilities akin to their younger peers, or their older peers, or perhaps somewhere in between. But how does one explain our finding that they fail to perform like either? A possible explanation rests in the *representational redescription* (RR) theory of development (Karmiloff-Smith, 1992). This theory, described in greater detail below, not only predicts a non-linear (or, what we will call, “U-shaped”) trajectory, it also pinpoints 18-months as an especially unique point in development.

**Distinguishing Representations from Behavior: An RR Model**

*Representational redescription* attempts to describe the process by which “implicit information *in* the mind subsequently becomes explicit knowledge *to* the mind” (Karmiloff-Smith, 1992:18). This representational change allows for theory creation and adaptation, which is a fundamental component to learning in childhood. It also suggests a number of predictions throughout development that ultimately allow for a more complete account of our data within the context of previous findings.

By way of background, RR is a model that is hypothesized to occur in three *phases.* Unlike a “stage” model of development (cf. Piaget, XXXX), these phases repeat multiple times within a single domain, and develop independently across multiple domains. This has two main implications: First, individual domains, and even single representations within domains, can develop on completely independent timescales. So, even if two domains share a concept X, for example, the manifestation of X in one domain may be far more complex, flexible, robust, efficient, rigid, etc. compared to the same concept X in another domain.

Relatedly, the second implication is that this account explains how cognitive weakness in one domain does not implicate a decline in functioning across *all* domains of cognition. This observation explains how children with autism, characterized by a decline in social cognition, still develop robust representations in domains concerned with number sense or language, for example. Piaget’s account of sweeping, incremental domain-specificity would be hard-pressed to explain the existence of domain-specific impairment, like autism or Williams Syndrome.

This characterization as a phase model also implicitly bears on the representational structure of the mind. Karmiloff-Smith’s (1992) ***account*** situates human cognition somewhere between prototypical accounts of domain-specificity and domain-generality. Specifically, she suggests that the *process* of RR is domain general, in that it occurs in every domain, but the *manifestation* of RR is nevertheless domain specific (i.e. *how* it occurs in each domain is unique to that particular domain; Karmiloff-Smith, 1992). Adopting a middle-ground between domain-specificity and domain-generality supports the claim that each domain undergoes specialization in its own regard, while simultaneously informing the specialization of other domains.

To illustrate the specific process of how information is specialize within a domain and also shared across domains, Karmiloff-Smith (1992) offers a specific characterization of each phase in the three-phase RR process. The first phase (P1) is characterized by a focus on external data. This “outsourcing” of information helps a particular domain achieve a state of “behavioral mastery,” or the consistently successful performance within that domain. Importantly, this performance can be identical to patterns of behavior displayed in the later phases without entailing precisely the same representations.

For example, when infants see what looks to be two portions of a rod shown behind an occluder, they form the same implicit expectations as adults—namely, that the two rod parts belong to a single, connected rod. This is evident in their longer looking time, interpreted as surprise, when the occluder is removed, and the rod segments are revealed to be spatially distinct, separated by a visible gap (Spelke, XXXX). While infants’ representation of the rod as being connected, and their subsequent behavior (surprise) when this expectation is violated, are implicitly identical to adults, their *explicit* knowledge differs vastly (e.g. an adult can tell you *why* they are surprised).

In the second phase of RR, emphasis is shifted away from implicit behaviors brought about by external data. Instead, focus is driven toward internally refining the currently-held representations. This often manifests in a decrease in behavioral performance relative to the first phase. One proposed explanation points to a developed lack of representational flexibility, due in part to the temporary suspension of input from other domains. Importantly, Karmiloff-Smith (1992) suggests that this decline in behavioral ability does not necessarily indicate degradation of the underlying representation. In a final act of reconciliation, external and internal data are integrated in the third and final phase of RR. This leads to behaviors similar to those in P1, but with representations that are far more flexible and robust.

In this way, the process of redescription largely maps to the U-shaped behavioral trajectory observed in measuring infants’ interpretations of rational action (*figure* 12). Specifically, younger infants display competence in a microdomain (i.e. rational action within the domain of agency, for example). Then as the infants grow older, performance in that domain appears to decrease (cf. our null result). Ultimately, though, infants recover their ability in a way that is more complex than originally observed (e.g. integrating knowledge of rationality *and* syntax).

Importantly, there is a distinction drawn between infants’ performance in *behavioral* tasks and the actual contents of their underlying *representations*. In P1, infants’ representations are *implicit* (Karmiloff-Smith, 1992). While this may produce accurate, or at least predictable behavioral patterns, the representations themselves are not yet fully developed or available to other domains of cognition. It is not until P3, when these representations become *explicit* knowledge to the infant, that they develop their flexibility across domains (Karmiloff-Smith, 1992). Thus, couching 18-month-olds’ understanding of rationality within the domain of language may have preceded their ability to manipulate these representations across domains, explaining the divergence in behavioral patterns.

Thus, if our data represent a felicitous snapshot of 18-month-olds’ development, this paints an interesting characterization of the behavioral trajectory from early infancy to toddlerhood. Crucially, though, our results do not bear on the actual constituents of the representations underlying the observed behaviors. The RR model predicts that these representations mature in a stable manner across development (Karmiloff-Smith, 1992), but confirming this would necessitate characterizing a pair of representations from a single domain at two distinct time-points, comparing them, and ideally proposing a mechanism by which one may arise from the other. This is an endeavor that is beyond the scope of the present study. However, future research may begin to address this pursuit by putting a methodological focus on measuring more *implicit* manifestations of a domain’s inner processes (e.g. by measuring implicit looking patterns rather than explicit imitation behaviors).

**Conclusion**

In a series of two experiments, we examined the relationship between infants’ judgments of rational, goal-directed action and their sensitivity to syntactically conveyed cues to meaning. This work was an extension of Kline and Snedeker (2015), and sought to construct a developmental trajectory for the onset of a mechanism that interprets both the behavioral and linguistic cues. Our findings suggested that 18-month-olds not only respond differently than their 24-month-old peers, but that they also do not produce the more primitive patterns predicted by an understanding of rational action, alone (cf. Gergely *et al.*, 2002). While this may be the result of the *hands occupied* condition failing to be sufficiently salient or plausible to the participants, an alternate account suggests that our null findings represent a divergence in behavioral patterns, but not representations, across the development of an integrated language-rationality mechanism.

However, this claim is unavoidably constrained by two factors. First, and most consequentially, it is not universally accepted that *representational redescription* is an accurate model of development. If this model is not a valid explanation of our results, then future research should attempt to parse whether our findings represent a confirmed failure, the weakness or context-dependency of our hypothesized effect, or unforeseen methodological errors. This also necessitates putting forth an account that explains the failure of 18-month-olds, especially in light of this non-linear developmental trajectory.

However, if we concede that the RR model is valid, the second constraint concerns the inability of our particular study to characterize the true nature of infants’ underlying representations. To that end, future work should attempt to: (1) identify a more detailed account of the relevant representations in the domains of both language and rational action; (2) characterize what a “joint” representation (explicitly present in P3) may look like; and (3) detail the process by which these representations are integrated.

Pursuing these future lines of research has implications beyond simply addressing whether infants are able to utilize cross-domain knowledge to construct an abstract representation of events in the world, and the language we use to describe them. By extension, we come closer to modeling the interconnectivity of the human mind and the domains of knowledge contained therein. We also begin to create a more accurate picture of the developmental origins and trajectory of these mechanisms. Ultimately, this may help to inform future research that attempts to understand both the nuanced micro-level processes of the mind, as well as the macro-level characterization of human cognition and behavior more broadly.

**References**

Beisert, M., Zmyj, N., Liepelt, R., Jung, F., Prinz, W., & Daum, M. M. (2012). Rethinking

‘rational imitation’in 14-month-old infants: a perceptual distraction approach. PLoS One, 7(3), e32563.

Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective imitation of in‐group

over out‐group members in 14‐month‐old infants. Child Development, 84(2), 422-428.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G Miller (Eds.)

*Linguistic theory and psychological reality* (pp. 264-293). Cambridge, MA: MIT Press.

Chen, M. L., & Waxman, S. R. (2013). “Shall we blick?”: Novel words highlight actors'

underlying intentions for 14-month-old infants. *Developmental psychology*, *49*(3), 426.

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short form

versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics*, *21*, 95–115.

Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation

in preverbal infants. Nature, 415(6873), 755-755. Chicago

Gleitman, L. (1990). The structural sources of verb meanings. Language acquisition, 1(1), 3-55.

Chicago

Hagedorn, J., Hailpern, J., & Karahalios, K. G. (2008, May). VCode and VData: illustrating a

new framework for supporting the video annotation workflow. In Proceedings of the working conference on Advanced visual interfaces (pp. 317-321). ACM.

Hirsh-Pasek, K., Gleitman, H., Gleitman, L. R., Golinkoff, R., & Naigles, L. (1988, October).

Syntactic bootstrapping: Evidence from comprehension. In 13th Annual Boston University Conference on Language Development.

Király, I. (2009). The effect of the model’s presence and of negative evidence on infants’

selective imitation. Journal of Experimental Child Psychology, 102(1), 14-25.

Király, I., Csibra, G., & Gergely, G. (2013). Beyond rational imitation: Learning arbitrary means

actions from communicative demonstrations. Journal of Experimental Child Psychology, 116(2), 471-486.

Kline, M. & Snedeker, J. (2015). I daxed my toy while blicking to it: 2-year-olds use syntax to

infer an actor’s intentions n a rational-action paradigm. Paper presented at the 37th Annual Conference of the Cognitive Science Society, Pasadena, CA.

Locke, J. (1948). An essay concerning human understanding, 1690. Chicago

Marcus, G. F. (1993). Negative evidence in language acquisition. *Cognition*, *46*(1), 53-85.

Meltzoff, A. N. (1988). Infant imitation after a 1-week delay: long-term memory for novel acts

and multiple stimuli. *Developmental Psychology*, *24*(4), 470.

Meltzoff, A. N. (1995). Understanding the intentions of others: re-enactment of intended acts by

18-month-old children. Developmental psychology, 31(5), 838.

Nielsen, M., Simcock, G., & Jenkins, L. (2008). The effect of social engagement on 24‐month‐

olds’ imitation from live and televised models. *Developmental science*, *11*(5), 722-731.

Paulus, M., Hunnius, S., Vissers, M., & Bekkering, H. (2011a). Bridging the gap between the

other and me: The functional role of motor resonance and action effects in infants’ imitation. *Developmental science*, *14*(4), 901-910.

Paulus, M., Hunnius, S., Vissers, M., & Bekkering, H. (2011b). Imitation in infancy: rational or

motor resonance?. Child development, 82(4), 1047-1057.

Phillips, A. T., & Wellman, H. M. (2005). Infants' understanding of object-directed action.

Cognition, 98(2), 137-155.

Pinker, S. (1994). How could a child use verb syntax to learn verb semantics?. Lingua, 92, 377-

410.

Southgate, V., Johnson, M. H., & Csibra, G. (2008). Infants attribute goals even to

biomechanically impossible actions. *Cognition*, *107*(3), 1059-1069.

Zmyj, N., Daum, M. M., & Aschersleben, G. (2009). The development of rational imitation in 9-

and 12-month-old infants. *Infancy*, *14*(1), 131-141.

Zmyj, N., Buttelmann, D., Carpenter, M., & Daum, M. M. (2010). The reliability of a model

influences 14-month-olds’ imitation. *Journal of Experimental Child Psychology*, *106*(4), 208-220.

Zmyj, N., Daum, M. M., Prinz, W., Nielsen, M., & Aschersleben, G. (2012). Fourteen‐Month‐

Olds' Imitation of Differently Aged Models. *Infant and Child Development*, *21*(3), 250-266.

Zmyj, N., & Buttelmann, D. (2014). An integrative model of rational imitation in infancy. Infant

Behavior and Development, 37(1), 21-28.