

A THING IS WHAT WE SAY IT IS:  
REFERENTIAL COMMUNICATION AND INDIRECT CATEGORY LEARNING

by

John Konstantinos Voiklis

Dissertation committee:

Professor James E. Corter, Sponsor  
Professor Deanna Kuhn  
Professor Robert O. McClintock, Chair  
Professor Peter T. Coleman  
Professor Oded Koenigsberg

Submitted in partial fulfillment of the  
requirements for the degree of Doctor of Philosophy  
under the Executive Committee of the Graduate  
School of Arts and Sciences  
COLUMBIA UNIVERSITY

2008

A Thing Is What We Say It Is:  
Referential Communication and Indirect Category Learning

John Konstantinos Voiklis

Submitted in partial fulfillment of the requirements for  
the degree of Doctor of Philosophy  
under the Executive Committee of The Graduate School of  
Arts and Sciences

COLUMBIA UNIVERSITY

2008

A Thing Is What We Say It Is:  
Referential Communication and Indirect Category Learning

John Konstantinos Voiklis

Professor James E. Corter, Sponsor

Submitted in partial fulfillment of the requirements for  
the degree of Doctor of Philosophy  
under the Executive Committee of The Graduate School of  
Arts and Sciences

COLUMBIA UNIVERSITY

2008

© 2008

John Voiklis  
All Right Reserved

## ABSTRACT

A THING IS WHAT WE SAY IT IS:

REFERENTIAL COMMUNICATION AND INDIRECT CATEGORY LEARNING

John Konstantinos Voiklis

This study investigates the interaction of referential communication and the structure of perceptual features on the joint processes of inventing a referential lexicon for novel objects and discovering the functional significance of those objects during an indirect category learning activity. During the learning task, participants worked either individually or as cooperative dyads to learn four combinations of orthogonal functional features—nutritive vs. not nutritive and destructive vs. not destructive—that defined four categories of fictional extra-terrestrial creatures. These categories were not specifically identified or labeled; rather, participants had to infer them indirectly as they predicted the functions. Also, these functionally defined categories exhibited a complex perceptual structure: a unidimensional (simple) rule predicted one function, while a family resemblance (complex) sub-structure predicted the other function. The function-learning task yielded function prediction data. In addition, each learner worked individually to sort the creatures (pre- and post-function learning) and to predict their functions in an individual function prediction posttest that also yielded selective attention data.

Together, the prediction data, sort data, and selective attention data supported three *a priori* hypotheses. Referential communication generates conceptual homogeneity (H3) and enhances indirect category learning (H1), though simple rules are learned earlier and better than complex relationships (H2). In explaining the learning advantages

observed among dyadic learners, I argue that referential communication may highlight attention to relationships between features (perceptual and functional) and actions as well as render such relationships more memorable. Moreover, communication may foster greater motivation among collaborators and may allow them to take advantage of the differing expectations and heuristics each collaborator brings to the task. In explaining the simplicity advantages observed among dyadic learners, I argue that referential communication may provide explicit “rules” for otherwise implicit (and perhaps more difficult) judgements. Dyads appear to have established reference to simple rules earlier than they established reference to complex rules; thus, they could explicitly (and perhaps more easily) learn the simple rule earlier than the complex rule. Finally, in explaining the conceptual homogeneity between and within dyads, I consider whether communication pushes “public” conceptualizations and publicly-formed “private” conceptualizations towards a limited range of widely shareable conceptual structures.

## TABLE OF CONTENTS

List of Tables.....	v
List of Figures.....	vi
Acknowledgements.....	ix
Dedication .....	x
I. Introduction.....	1
Background.....	1
Joint Activity & Referential Communication.....	1
An Example of Coordination Through Referential Communication.....	1
“Public” Categorization.....	2
Shareability .....	3
Structural Constraints.....	4
Purpose.....	5
Structure of the Dissertation.....	6
II. Literature review .....	8
Plausible Constraints on Human Category Learning.....	8
Structural Constraints.....	9
Informational & Biological Constraints.....	10
Structural & Cognitive Constraints.....	12
Structure, Cognition, and Activity .....	13
Person-Related Constraints.....	14

Language Constraints .....	15
Category labels as conceptual cues .....	16
Category labels as conceptual manipulatives .....	16
Category labels as conceptual manipulators.....	17
Summary of Plausible Constraints.....	18
Conceptual Coordination Via the Processes of Referential Communication ...	19
Referential Games.....	19
Coordination of Attention and Intention in Referential Games.....	20
From Private to Public Categories.....	20
Joint Attention, Joint Intention, and Joint Reference .....	21
From Joint Reference to Referential Convention .....	22
From Public to Private Categories .....	23
Conclusion and Hypotheses .....	25
III. method.....	28
Purpose & Overview.....	28
Participants .....	29
Design .....	29
Materials.....	31
The Design of the Category Structure.....	31
The Design of the Stimuli.....	32
Tasks .....	33
The Function Prediction Task.....	33
The Sorting Tasks .....	35



The Attention Allocation Task.....	36
Procedure.....	36
Data Analysis.....	37
Deriving measures of prediction accuracy .....	37
Deriving measures of attention allocation .....	38
Deriving types of sorting explanations & measures of structural similarity .....	39
IV. Results.....	41
Referential communication enhances indirect category learning.....	41
Functional-category prediction accuracy.....	41
Posttest category prediction accuracy and attention allocation .....	44
The structure of post-learning sort clusters .....	46
Referential communication generates conceptual homogeneity .....	47
Within-Dyad and Between-Dyad Sorting Homogeneity.....	47
Within-Dyad Referential Homogeneity .....	49
Referential communication enhances simple-rule learning more than complex- rule learning.....	51
Function prediction accuracy .....	51
Posttest function prediction accuracy and attention allocation.....	52
Structural similarity of post-learning sort clusters to the rule-defined clusters .....	55
V. Discussion.....	57
Referential communication enhances indirect category learning.....	58
Summary of Results .....	58

Explanation of dyadic advantage .....	59
Implications of dyadic advantage.....	63
Further exploration of dyadic advantage.....	64
Referential communication enhances simple-rule learning more than complex- rule learning.....	65
Summary of Results .....	65
Explanation of simplicity advantage.....	65
Implications of simplicity advantage .....	67
Further exploration of simplicity advantage.....	68
Referential communication generates conceptual homogeneity .....	69
Summary of Results .....	69
Explanation of conceptual homogeneity .....	69
Implications of conceptual homogeneity.....	71
Further exploration of conceptual homogeneity.....	71
References .....	73
Appendix A .....	83
Descriptive and Point-Based Feedback.....	83
Appendix B.....	85
Examples of Stimuli.....	85
Appendix C.....	86
Instructions for Tasks and Task Interfaces.....	86

## LIST OF TABLES

Table 1. Category Structure represented in binary notation. ....	31
Table A1. Descriptive Feedback for the Function Prediction and Attention Allocation Tasks. ....	83
Table A2. <i>Energy Meter</i> : Rewards and Penalties for the Function Prediction and Attention Allocation Tasks. ....	84

## LIST OF FIGURES

Figure 1. Compares the rate of functional-category prediction accuracy of dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials). Error bars indicate confidence interval. ....	42
Figure 2. Compares the rate of functional-category prediction <i>accuracy</i> of dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) during the Attention Allocation task. Error bars indicate confidence interval. ....	43
Figure 3. Compares the average number of features per stimulus that dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) uncovered during the Attention Allocation task. Error bars indicate confidence interval.....	44
Figure 4. Compares the <i>structural similarity</i> of the post-learning sort clusters produced by dyadic learners and individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) to the “true” category clusters. Error bars indicate confidence interval. ....	45
Figure 5. Compares the probability of citing a category (whether functionally-defined or behaviorally-defined) when dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning	

trials) explained their post-learning sort clusters. Error bars indicate confidence interval. ....	46
Figure 6. Compares the mean pair-wise similarity of the post-learning sort clusters produced by actual dyads vs. between-dyad pairings and pairs of individual learners. Error bars indicate confidence interval. ....	48
Figure 7 A-C. Compare the <i>simple-rule</i> function prediction accuracy rates vs. <i>family-resemblance</i> function prediction accuracy rates of dyadic learners vs. individual learners who completed 160 learning trials vs. individual learners who completed 320 learning trials. Error bars indicate confidence interval. ....	50
Figure 8. Compares the <i>simple-rule</i> function prediction accuracy rates vs. <i>simple-rule</i> function prediction accuracy rates of dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) during the Attention Allocation task. Error bars indicate confidence interval. ....	53
Figure 9. Compares the probability that a dyadic learner vs. individual learner (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) would uncover a <i>family-resemblance</i> related versus <i>simple-rule</i> related feature on any Attention Allocation task trial. Error bars indicate confidence interval. ....	54
Figure 10. Compares the <i>structural similarity</i> of the post-learning sort clusters produced by dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) to the category	

clusters defined by one or another <i>type of rule</i> . Error bars indicate confidence interval. ....	55
Figure 11. Compares the probability that dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) mentioned features relating to one or another type of rule when explaining their post-learning sorts. Error bars indicate confidence interval. ....	56
Figure B1. Examples of the binary feature values that instantiate the stimuli and the perceptual structure of the functional categories.....	85
Figure C1. Describes the scenario for the game tasks and provides instruction on how to execute function-prediction related actions. ....	86
Figure C2 A & B. Spotter's Function Prediction Task Interface. A (top) = Single Player, and B (bottom) = Two Player.....	87
Figure C3 A & B. Function Prediction Task Instructions. A (top) = Single Player, and B (bottom) = Two Player.....	88
Figure C4. Two-Player Beamer Function Prediction (FP) Task Interface. ....	89
Figure C5. Interface for the PRE-sort and POST-sort tasks. ....	90
Figure C6 A & B. Instructions for the PRE-sort (top) and POST-sort tasks (bottom). ....	91
Figure C7. Attention Allocation Task Interface.....	92
Figure C8. Instructions for Attention Allocation Task, including how to execute task-related action. ....	93

## ACKNOWLEDGEMENTS

First, I need to thank my advisor and collaborator, Professor James Corter, whose generous contribution of time, energy, and intellect helped me bring this project to fruition. I look forward to our continued collaboration.

Also, I thank the members of my dissertation committee—Professors Robert McClintock, Deanna Kuhn, Peter Coleman, and Oded Koenigsberg—who as teachers and interlocutors variously helped me to prepare for my current and future endeavors.

I must acknowledge the support of the Ben and Grace Wood Fellowship, which, for the first three years of my graduate school *career*, afforded me the freedom to pursue learning beyond the confines of any curriculum and research beyond the confines of any disciplinary commitments.

In the same breath, I must thank Manu Kapur (now Professor Kapur) and Professor Charles Kinzer; our research collaboration and our coffee klatch contributed immeasurably to my development as a thinker and as a scientist.

I must also thank Professor John Black, who believed that a poet, erstwhile hacker, and dabbler in artificial life and artificial intelligence had something to contribute to the cognitive study of education and learning technologies.

Finally, I must thank all those who participated in my study for the seriousness and vigor with which they engaged the experimental tasks.

## **DEDICATION**

For my parents, Asimina and Konstantinos, who always expected me to aspire to ever greater accomplishments. For my children, Kosta and Magda, whose curiosity and greed for learning I aspire to emulate. For Charlotte, my wife, who, whether or not she shares in every one of my aspirations, gladly shares in the burden of fulfilling any and all of them.



## I. INTRODUCTION

### Background

#### *Joint Activity & Referential Communication*

Human beings engage in myriad joint activities: a parent and child jointly build a Lego robot; two families jointly plan and execute a wedding celebration; a far-flung group of scholars and researchers jointly develop a domain of knowledge. The successful performance of these and other joint activities requires the coordination of actions (Schelling, 1960); the coordination of actions presupposes the coordination of intentions, assumptions, and beliefs that drive those actions (Lewis, 1969; Stalnaker, 1978, 2002; Clark, 1996). Referential communication can facilitate these multiple levels of coordination (*cf.*, Bangerter & Clark, 2003; Sacks, Schegloff, & Jefferson, 1974).

#### *An Example of Coordination Through Referential Communication*

Imagine the following exchange between Alpha and Beta (named for their status in the kitchen) as they prepare a meal to impress their new girlfriends, Bonita and Belle (named for their beauty). Inspired by a mutton shank, Alpha requests, “heat up the tagine, while I dice the aromatics.” At the cookware cabinet, Beta asks, “the heavy one with the flat lid?” “No,” clarifies Alpha, “that’s the dutch oven. We need the ceramic one with the conical lid.” “Ok,” Beta confirms and asks, “should I put on a kettle for couscous, while

I'm at it?" While the conversation between Alpha and Beta could continue in this way over multiple pages, this much of the exchange suffices to illustrate how referential communication helps them coordinate their culinary seduction of Bonita and Belle.

### ***"Public" Categorization***

Referential communication serves as a form of externalized or "public" cognition (*cf.*, Wittgenstein, 2001 [1958]), including "public" processes of categorization (see Russell, 1905; Brown, 1958; Grice, 1975; Cruse, 1977; Barr & Kronmüller, 2006 for various formulations of the conceptual function of reference). In the scenario, Alpha performs three such "public" categorizations: "aromatics" are a category of vegetables—onions, carrots, celery, *et al.*—used to flavor a dish, while "tagine" and "dutch oven" are subcategories of braisers—shallow, tightly-lidded pots used for slowly cooking a dish in its own condensation. By using a particular reference at a particular level of reference, communicating actors *direct each other's attention* (Kronmüller & Barr, 2007; Metzing & Brennan, 2003) to those features that allow each of them *to differentiate the target referent* from other possible referents (E.V. Clark, 1987; Mervis & Crisafi, 1982; Murphy & Brownell, 1985). Alpha instructs Beta to "heat up the tagine" (as opposed to "heat up the pot" or "heat up the braiser") in order to help Beta differentiate the target cookware from the stock pots, roasting pans, and skillets in the cookware cabinet. Beta understands that a tagine is kind of braiser, but needs Alpha to direct his attention to the tagine's differentiating features: ceramic *vs.* cast iron, conical lid *vs.* flat lid. Further, the use of a particular reference at a particular level of reference can *direct joint attention* to those features that allow each actor *to infer the referent's significance* to the activity (Brown,

1958; Corter & Gluck, 1992). The tagine originates from Morocco; so, Beta infers from that feature that Alpha has proposed that they prepare a Moroccan dish and *confirms* that joint intention by asking if he should boil water for couscous. Beta's confirmatory question completes the "public" categorization of the tagine. The successful use of a particular reference at a particular level of reference should elicit *an action that confirms the joint construal* of both the target and its significance to the activity (Austin, 1976; Krauss & Weinheimer, 1966; Wilkes-Gibbs & Clark, 1992; Clark & Schaefer, 1989).

### *Shareability*

Cooking is a recurring, often institutionalized, coordination problem; therefore Alpha and Beta can rely heavily on conventions of referential communication—a repertoire of mutually-known, mutually-salient, and mutually-expected associations between reference and concept. Adherence to referential conventions can minimize the cognitive effort of directing *joint* attention, confirming *joint* construal, and executing *joint* intentions (Clark, 1996, Lewis, 1969). In this way, referential conventions convey conceptual information in a highly *shareable* form (*cf.*, Freyd, 1983). Nevertheless, when faced with a new or unfamiliar activity, communicating actors often rely on *ad hoc* referential conventions (Garrod & Anderson, 1987; Brennan & Clark, 1996; *et al.*).

Imagine Alpha and Beta repairing the electric ignitor on their stove; unfamiliar with the conventional names for various circuit parts, Beta points out, "that mini-bulb has burned out," referring to a glass cartridge fuse. Such *ad hoc* conventions also reduce the cognitive effort of directing joint attention (Clark Schreuder, & Buttrick, 1983; Clark, 1972; Clark & Marshall, 1981), confirming joint construal (Wilkes-Gibbs & Clark, 1992;

Isaacs & Clark, 1987; Clark & Krych, 2004), and executing joint intentions (Clark & Lucy, 1975; Francik & Clark, 1985). The use of “mini-bulb” directs Alpha’s attention to a small glass object, containing a fine metal element. Alpha responds, “yeah, the filament has melted,” confirming his joint construal of “mini-bulb.” Moreover, these *ad hoc* conventions may affect how each actor later sorts conventionally named objects (Markman & Makin, 1998) and how each actor later judges the similarity and/or typicality of objects to a conventionally named category (Malt & Sloman, 2004). For example, one might expect Alpha or Beta to search for glass cartridge fuses in the lighting aisle of the hardware store rather than the circuitry aisle.

Nevertheless, *ad hoc* referential conventions can vary in shareability. The *mini-bulb* category ultimately fails. It does not enable either Alpha or Beta *to infer the functional significance* of the fuse—a device for protecting circuits from the power surges that today burned a fuse and tomorrow might burn their building. For Alpha and Beta this failure will not persist; the hardware clerk will likely provide the necessary knowledge. In an entirely novel activity, communicating actors lack both prior knowledge and access to third-party experts. They must invent a referential lexicon for the objects, actions, and events in the activity environment as they discover the differentiating and significant features of those objects, actions, and events. Under such conditions, shareability may vary with more basic factors, such as feature structure.

### ***Structural Constraints***

When engaged in *private categorization*—category learning during individual activity—human beings tend to learn simple relationships between features and their

significance more easily than complex relationships (*cf.*, Feldman, 2003b). Imagine Bonita, a bee researcher, trying to diagnose colony collapse. She easily diagnoses parasite infestation, based on the presence or absence of mites; diagnosing vanishing bee syndrome, with its varying constellation of symptoms and uncertain causes, requires more effort and yields tentative results. Then again, simply labeling a privately learned category can enhance an individual's ability to infer the category's significance to the activity despite fairly complex, even contradictory, featural information (Yamauchi & Markman, 2000). Even a tentative diagnosis of a "syndrome" should prompt Bonita to suspect factors that cause general disruptions in the immune systems of the bees. What happens, though, when Bonita invites Belle, a designer of agent-based models, to collaborate on a simulation of colony collapse; will their conversation bog down in the complexities of vanishing bee syndrome, or will their emerging convention of referential labels enhance their joint reasoning? How structural complexity affects the shareability of *ad hoc* conventions remains uncertain.

### **Purpose**

With the present study, I investigated the interaction of referential communication and the structure of perceptual features on the joint processes of inventing a referential lexicon for novel objects and discovering the functional significance of those objects during an indirect category learning activity. To that end, participants worked either

individually or as cooperative<sup>1</sup> dyads to learn four combinations of orthogonal functional features—nutritive vs. not nutritive and destructive vs. not destructive—that defined four categories of fictional extra-terrestrial creatures. These categories were not specifically identified or labeled; rather, participants had to infer them *indirectly* as they predicted the functions. Also, these functionally defined categories exhibited a complex perceptual structure: a unidimensional (simple) rule predicted one function, while a family resemblance (complex) sub-structure predicted the other function. This function-learning task yielded function prediction data. In addition to the main function-learning task, each learner worked individually to sort the creatures (pre- and post-function learning) and to predict their functions in an individual function prediction posttest that also yielded selective attention data. Together, the prediction data, sort data, and selective attention data demonstrated the overall affects of referential communication on the extra-linguistic aspects of concept learning and the differing affects of communication on the learning of simple versus complex relationships between perceptual and latent features.

### **Structure of the Dissertation**

What follows is a detailed report of this investigation, organized into four chapters. In the *Literature Review* chapter, I attempt to integrate findings and constructs from two research traditions: cognitive research on category learning, both in general and with an explicit focus on lexically labeled categories, and psycho-linguistic research on conversation and referential communication. From this integrated review, I glean three

---

<sup>1</sup> Dyads had positively *interdependent* goals: the success of each actor depended on the other (Deutsch, 1949; Johnson & Johnson, 1989).

hypotheses: referential communication generates conceptual homogeneity (H3) and enhances indirect category learning (H1), though simple rules are better learned than complex relationships (H2). The *Method* chapter provides details on the experimental design that I used to test these hypotheses, as well as on how and why I collected and analyzed the various data. Next, in the *Results* chapter, I highlight how the learning data and the linguistic data supported the hypotheses. Finally, I interpret the major findings in the *Discussion* chapter. In particular, I focus on how referential communication directs attention to relationships between features (perceptual and functional) and action. In addition, I argue that referring expressions may render such relationships more memorable and may provide explicit “rules” for otherwise implicit judgements. Finally, I speculate on whether communication pushes “public” conceptualizations and publicly-formed “private” conceptualizations towards a limited range of widely shareable conceptual structures.

## II. LITERATURE REVIEW

When faced with a new or unfamiliar joint activity, communicating actors quickly converge on *ad hoc* referential conventions (Garrod & Anderson, 1987; Brennan & Clark, 1996; *et al.*). These *ad hoc* conventions entail a system of shared or “public” categories of the objects, actions, and events in the activity environment. The *shareability* of a convention derives from the extent to which it minimizes the joint cognitive effort of sharing attention and intentions towards those objects, actions, and events (*cf.*, Freyd, 1983). In the *wild*, shareable conventions emerge from the coupled evolution of both the linguistic and extra-linguistic aspects of the various “public” categories during conversationally driven joint activity. I consider this evolution from the perspective of two research traditions: cognitive research on “private” category learning, both in general and with an explicit focus on lexically labeled categories; and psycho-linguistic research on conversation and referential communication. Integrating these traditions yields hypotheses about the “public” category learning entailed in the emergence of shareable conventions that one could not derive from either tradition alone.

### Plausible Constraints on Human Category Learning

While conventions such as driving on the left-hand versus right-hand side of the road may seem arbitrary, one rarely finds an entirely arbitrary referential convention. As a system of “public” categorizations, a referential convention is subject to constraints on



the human conceptual system. Some constraints originate from outside human beings. For example, both the natural and artificial worlds exhibit structural regularities (*e.g.*, Berlin, Breedlove, & Raven, 1966; Simon, 1956). Other constraints originate from within human beings. These include, limits on memory and selective attention (*e.g.*, Shepard, Hovland, and Jenkins, 1961), the limits of embodiment (*e.g.*, Gallese & Lakoff, 2005), and the limits of personal experience (*e.g.*, Murphy & Medin, 1985), among others.

### ***Structural Constraints***

Among the various constraints addressed by the categorization and decision-making literature, the structure of the features defining categories of objects, actions, and events appears to exert the most pervasive or, at least, most discernible influence on category learning and use (Rosch, Mervis, Gray, Johnson, & Braem, 1976; Corter & Gluck, 1992; Simon, 1956; Todd & Gigerenzer, 2007). For example, population-wide regularities in how individuals name and classify various natural kinds, including colors (*e.g.*, Kay *et al.*, 2007 [2003]), kinship (*e.g.*, Goodenough, 1965; Romney & D'Andrade, 1964; Romney, Boyd, Moore, Batchelder, & Brazill, 1996), as well as plants and animals (*e.g.*, Berlin, Breadlove & Raven, 1973; Diamond, 1966; Bulmer, 1967; Hunn, 1977), appear to reflect statistical regularities in the environment (Berlin *et al.*, 1966; Rosch *et al.*, 1976). This concordance holds even after one accounts for the arbitrary distinctions within any particular population-wide convention (Malt, Sloman, & Gennari, 2003a, 2003b; Malt, Sloman, Gennari, Shi, & Wang, 1999).

Moreover, individuals tend to name and identify categories of objects (*e.g.*, Rosch *et al.*, 1976; Jolicoeur, Gluck, & Kosslyn, 1984; Tversky & Hemenway, 1984; Murphy &

Smith, 1982), actions (Tomasello & Merriman, 1995; Vallacher & Wegner, 1987), and events (Morris & Murphy, 1990; Rifkin, 1985) at a level of abstraction—the so-called basic level—that permits the average person to infer the maximal number of differentiating features of those objects, actions, and events (Gluck & Corter, 1985). For example, if Alpha, the imaginary character from the *Introduction*, were talking to Bonita, a culinary novice, he would likely refer to the tagine as “the *pot* with the conical lid” instead of using the category label *tagine* or even *braiser*. Similar tendencies in the naming and identification of experimenter-designed categories suggest that statistical regularities in the environment may push human conception towards the most informative level (Cortier, Gluck, & Bower, 1988; Murphy & Smith, 1982; Hoffmann & Ziessler, 1983). In fact, various analytical models (*e.g.*, Gosselin & Schyns, 2001; Gluck & Corter, 1985; Rosch *et al.*, 1976; Jones, 1983) can approximate the observed primacy of basic-level categorization as a function of featural information.

### ***Informational & Biological Constraints***

Phenomena like basic-level primacy appears so pervasive that one might speculate whether biological evolution has *hard-wired* certain conceptual structures (*e.g.*, Atran, 2005; Medin & Atran, 2004; Shepard, 1992; Tooby & Cosmides, 1989; Rosch, 1975). By this account, the tendencies to name and identify both natural and artificial kinds has evolved as an adaptive response to fairly persistent informational structures in the environment (Atran, 2005), and any observed deviations from such “universal” tendencies represents a devolution of the conceptual system (Atran, Medin, & Ross, 2004). To a great extent, the adaptionist account relies on discontinuities between human

beings and other primates (Hauser, 2005). Nevertheless, many primate species exhibit homologous understandings of various concepts, such as small quantities (Dehaene, 2001; Hauser, 2000; Hauser *et al.*, 2000, 1996), gravity (Hood, Hauser, Anderson, & Santos, 1999), distinctions between living and artifactual kinds (Hauser, 1997), and functional distinctions among artifactual kinds (*ibid.*). Moreover, non-human primates trained to use arbitrary symbols can make abstract judgements that are seldom recognized in the wild (Thompson, Oden, & Boysen, 1997).

For the most part, the differences between the human and non-human conceptual systems appear quantitative rather than qualitative (Deacon, 2000). A coupling of a quantitatively different conceptual system with quantitatively different capacities for opportunistic learning suffices to produce qualitatively different results (Wagner & Wagner, 2003). Add to that, a quantitatively different capacity for sharing mental states (*e.g.*, Meltzoff & Andrew, 2007; Saxe, 2006; Tomasello, Carpenter, Call, Behne, & Moll, 2005; Bloom, 2002) via *heritable* media (*e.g.*, language and artifacts; *cf.*; Hutchins & Hazelhurst, 1992; Cavalli-Sforza & Feldman, 1983; Vygotsky, 1978), and strongly adaptionist accounts of the human conceptual system appear, at best, unnecessary (Gould & Lewontin, 1979).

At most, one might argue that comparable sensori-motor capacities among human beings may yield comparable experiences and, thus, comparable conceptions of a common environment (Gallese & Lakoff, 2005; Pfeifer & Bongard, 2006). For instance, the particular configuration of the human body yields a particular experience of physical space, which may, in turn, yield a spatially-analogous structure to both concrete and abstract concepts (Moyer, 1973; Paivio, 1978; Tversky, 2005).

### *Structural & Cognitive Constraints*

Cognitive constraints, such as limits on working memory and selective attention, represent a less controversial expression of biology. Structural and cognitive constraints appear to intermingle in the oft-cited study by Shepard, Hovland, and Jenkins (1961), where varying the number and combination of diagnostic dimensions of artificial category structures varied the time and effort required to learn those structures (see also Nosofsky, Gluck, Palmieri, Mckinley, & Glauthier, 1994; Love, 2002). Assuming that category learners attempt to optimize selective attention across diagnostic dimensions, one could attribute the increasing cognitive effort directly to increasing demands on selective attention and working memory (*e.g.*, Nosofsky *et al.*, 1994). More often, though, category learners rely on satisficing solutions (Simon, 1955), inferring categories from the simplest description derivable from known exemplars (Feldman, 2003b; Chater & Vitányi, 2003). For example, in the *Introduction*, Beta relies on the diagnostic features of the familiar category *light bulb*—glass object containing a filament—to identify the unfamiliar *fuse* as a “mini bulb.” This sort of “fast and frugal” reasoning conserves cognitive resources (*e.g.*, Todd & Gigerenzer, 2001; Gigerenzer & Goldstein, 1996) and has been observed in categorical decision-making (Matsuka & Corter, In Press; Medin *et al.*, 1987).

That said, the extent of one’s “frugality” depends on the compressibility of the category structure (Feldman, 2000, 2003a, 2006). For example, Shepard *et al.* (1961) designed each of their six category structures using three binary dimensions, or twelve bits of information. Expressed as a Boolean equation, the category structure with one diagnostic dimension and two non-diagnostic dimensions compresses to one bit, while

the category structure with three diagnostic and nonlinearly separable dimensions compresses no smaller than ten bits of information. In other words, the simplest heuristic for learning the latter category structure would require nothing short of memorizing each category exemplar (*e.g.*, Allen & Brooks, 1991). In the *Introduction*, Bonita's difficulty in diagnosing Vanishing Bee Syndrome attests to the excessive cognitive effort required in learning and using an incompressible category structure.

### ***Structure, Cognition, and Activity***

The increasing effort required for learning increasingly complex structures may stem, in part, from the particular activity—classification—that dominates laboratory studies on categorization (for review, see Markman & Ross, 2003). The interplay between structural, cognitive, and activity-related constraints becomes obvious when one contrasts alternative uses of categories (*ibid.*)—*e.g.*, classifying an object versus inferring its uncertain features. Whereas classification focuses attention on between-category featural information, those engaged in feature inference appear to allocate attention to within-category featural information, especially to specific feature values and the within-category correlations among these values (Yamauchi & Markman, 1998, 2000a, 2000b). For example, if one needed to classify a particular brazier as either a *tagine* or a *dutch oven*, one would attend to the category differentiating features—ceramic *vs.* cast iron, conical lid *vs.* flat lid. Alternatively, if one needed to infer whether a particular pot was appropriate for braising a mutton shank, one would attend to the width, depth, and thickness of the pot as well as how tightly the lid fits. Those engaged in a classification task tend to learn simpler, more compressible, category structures more easily than

complex category structures (*ibid.*). Those engaged in an inference task learn both simple and complex category structures with comparable effort (*ibid.*).

Activities that involve the indirect learning of categories—*e.g.*, looking for patterns among stimuli or rating their pleasantness—appear similar to inference tasks in attentional strategy and the effort expended in learning simple versus complex category structures (Love, 2002, 2003). Minda & Ross (2004) serves as a noteworthy example of indirect category learning. Participants predicted the food allotment for a sequence of imaginary animals. Successful predictions required the learning of a complex category structure, where both a unidimensional rule and family resemblance (multidimensional rule) predicted the food allotment. While participants could rely on either or both rules, attentional strategies varied with whether or not a classification task (direct category learning) preceded the prediction task (indirect category learning). Those who classified animals before predicting food allotments relied on the simple rule; those who only predicted food allotment distributed their attention across multiple dimensions (see also Ross, 1997, 1999).

### ***Person-Related Constraints***

Other constraints derive from the person (Murphy & Medin, 1985), including his or her prior knowledge (*e.g.*, Chi, Feltovich, & Glaser, 1981; Gauthier, Williams, Tarr, & Tanaka, 1998; Tanaka & Taylor, 1991; Murphy & Wright, 1984; Schvaneveldt *et al.*, 1985), current goals (Barsalou, 1983; Ratneshwar, Barsalou, Pechmann, & Moore, 2001), and the situational heuristics (*e.g.*, Gluck, Shohamy, & Myers, 2002) that connect prior knowledge and current goals. These individual differences may elicit varying prior

expectations of how the objects, actions, and events of a new, rare, or unfamiliar activity relate to one another (*c.f.*, Murphy & Medin, 1985). Moreover in trying to induce the *actual* relationships among the objects, actions, and events, differing individuals might use different heuristics (Lin & Murphy, 1997). In this way, person-related constraints can enhance or impede the learning of novel categories. For example, when Bonita and Belle build their computational simulation of Vanishing Bee Syndrome, they are likely to model complex interactions among variables like microwaves and genetically modified crops based on their prior expectations of a *syndrome*. Likewise, they are likely to exclude simple factors like mite infestation, which might account for many symptoms of the syndrome. Usually, when faced with contrary evidence, the category learner abandons misapplied prior hypotheses (Livingston & Andrews, 1995). The absence of glass cartridge fuses in the lighting aisle of the hardware store will prod Alpha and Beta towards the circuitry aisle. When prior hypotheses bear some resemblance to the observed evidence, though, mistaken hypotheses can persist and impede learning (*ibid.*). A simulation that includes microwaves and genetically modified crops but excludes mite infestation will not help beekeepers.

### ***Language Constraints***

What a category is called—the category label—and whether or not it is called anything at all can affect the learning and use of a category. Such effects fall short of the notion that language *is* thought (Davidson, 1975) or that language *determines* thought (Whorf, 1956), but goes far beyond the notion that category labels serve little purpose beyond that of another diagnostic feature among other diagnostic features (Anderson,

1991). Category labels appear to serve three interrelated functions: (1) as conceptual cues, (2) as conceptual *manipulatives*, (3) and as conceptual *manipulators*.

### ***Category labels as conceptual cues***

By the principle of contrast (E.V. Clark, 1987), different category labels signal different concepts. For example, labeling objects by different names can help children individuate those objects (Xu, 2002) and often leads them to look for differences among differently-labeled objects (Katz, 1963; Landau & Shipley, 2001) and for similarities among objects with the same name (Loewenstein & Gentner, 2005; Smith, Jones, & Landau, 1996; Waxman & Markow, 1995). Similarly, adults learn event categories better when verbs and/or syntax covaries with the events (Cabrera & Billman, 1996). Also, children often treat unknown labels as category labels for unknown objects and as feature labels for known objects (Markman & Wachtel, 1988). Finally, the mere presence of category labels can cue the category learner to look for meaningful patterns in what he or she perceives. Children tend to pay more attention to labeled categories than to unlabeled categories (Balaban & Waxman, 1997; Waxman & Booth, 2001) and adults learn labeled categories more quickly than unlabeled categories (Lupyan, Rakison, & McClelland, 2007). In all, category labels appear to make abstractions concrete and implicit judgements explicit (A. Clark, 2006, Vygotsky, 1986 [1962]).

### ***Category labels as conceptual manipulatives***

The concreteness of category labels reduces the cognitive effort of reasoning from and about abstract and complex concepts, much like Cuisenaire rods reduce the effort of learning and using mathematical concepts. Among the examples of the cueing function of



labels, category learners may have offloaded burdensome feature comparisons onto the category labels between which differences and similarities are easy to discern (A. Clark & Karmiloff-Smith, 1993). Beyond differentiation, children rely more on category labels than perceptual similarity when inferring unknown features (Gelman & Markman, 1986), and use labels to simplify relational judgements (Kotovsky & Gentner, 1996). Moreover, labels anchor abstract concepts like large (Gordon, 2004) and exact numerosities (Gellman & Gallistel, 2004) as well as conceptual manipulations like exact mental arithmetic (Beller & Bender, 2008; Pica, Lemer, Izard, & Dehaene, 2004). Finally, labeling categories enhances visual search by reducing the search space to the category-relevant features (Lupyan, 2008), or, put another way, by maximizing attention to diagnostic features.

### *Category labels as conceptual manipulators*

This attentional control points to a third function of category labels: manipulating the attentional, perceptual, and memory-related processes of category learning and use (*cf.*, Vygotsky, 1986 [1962], on how language scaffolds thought). For example, children appear to use the names of previously encountered objects to tune their attention to naming-relevant features of subsequently encountered objects (Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002). When adults classify objects using basic-level category labels they remember and, perhaps, perceive those objects as more similar to the prototype of the labeled category (Lupyan, 2008; see also Carmichael, Hogan & Walter, 1932; Daniel, 1972). Similarly, assigning descriptive labels to novel stimuli improves the sequential matching of those stimuli, a task heretofore considered exclusively visual. In contrast, describing the particular features of a particular face (Schooler and Engstler-

Schooler, 1990) or a particular wine (Melcher & Schooler, 1996) can impair later recognition of that face or that wine. Surprisingly, describing a face in terms of prototypical features does not distort memory (MacLin, 2002). In all, category labels appear to manipulate the conceptual system via a top-down activation of category features that attracts attentional processes and colors or occludes memory and perception.

### ***Summary of Plausible Constraints***

Structure, complexity, language, and activity (along with embodiment and person-related constraints) represent an *A-list* rather than a catalog of all plausible extra-conversational constraints on referential convention. These *A-list* constraints appear obvious constituents of any overarching shareability constraint, each may affect the joint cognitive effort of sharing a concept. Nevertheless, few empirical researchers have examined how referential communication might amplify, dampen, or distort these constraints (for exceptions see Markman & Makin, 1998, and Malt & Sloman, 2004). Intuitively, one might suspect that structural regularities in the world and the complexity of those structures should constrain *communicating* individuals in ways similar to *isolated* individuals, but how and to what extent is not certain. Moreover, many of the studies on use and usefulness of category labels assume that one can decouple language from communication. Language is for communicating, and communicating is for thinking jointly and in public (*cf.*, Wittgenstein, 2001 [1958]). The use of language requires both a *self* and an *other*—whether actual or implied.

## Conceptual Coordination Via the Processes of Referential Communication

### *Referential Games*

In new or rare or unfamiliar joint activities, actors often coordinate their actions through conversation (Clark, 1996). Since Wittgenstein (2001 [1958]), the processes of conversation, in general, and referential communication, in particular, have often been formulated as a game (cf., Higgins, 1981; Crawford & Sobel, 1982; Blume, Dejong, Kim, & Sprinkle, 1998; Pietarinen, 2006). The players, as speaker and addressee, set the rules of play and keep score through their ongoing joint construal of the various referents and their significance to the activity (Lewis, 1979). Coordination proceeds through a turn-taking process where, at any iteration, a speaker will refer to some object, action or event in the activity environment. The speaker's use of a particular reference at a particular level of reference functions as a *public categorization* of the referent, differentiating the referent and highlighting its significance to the activity (see Russell, 1905; Brown, 1958; Grice, 1975; Cruse, 1977; Barr & Kronmüller, 2006 for various formulations of the conceptual function of reference). In reply, the addressee can ratify the proposed categorization, seek clarification, or offer a counter-proposal (cf., Clark & Wilkes-Gibbs, 1986; Sacks *et al.*, 1974; Clark & Krych, 2004; Hulstijn & Maudet, 2006; W. Deutsch & Pechmann, 1982). The conversational turn continues until joint-construal is confirmed.

### ***Coordination of Attention and Intention in Referential Games***

To illustrate such a game, imagine Bonita and Belle collaborating in what for them is a novel activity, cooking. Rummaging in the pantry for ingredients, Bonita requests, “could you sharpen the big knife?” Grasping a carving knife, Belle asks, “this long one?” “Is it big?” Bonita checks, to which Belle insists, “long is big.” “Fine, the wide one,” clarifies Bonita. Carrying shallots from the pantry, Bonita finds a sharpened cleaver on her cutting board and responds, “well, *you* can use that to hack apart the rib chops, but that’s not a dicing tool.” “It’s wide,” persists Belle. “I need the one that widens from a point to a round belly,” Bonita further clarifies. “Oh, this one,” pulling the chef’s knife from the block; “it’s like half a bow” she counter proposes. “Yes,” Bonita confirms, “the half-bow knife.”

### ***From Private to Public Categories...***

Early in conversation, each actor is likely to propose and/or counter-propose relatively idiosyncratic or private categorizations of the objects, actions, and event in the activity environment (Horton & Keysar, 1996; Keysar, Barr, Balin, & Brauner, 2000; Keysar, Barr, Balin, & Paek, 1998; Kronmüller & Barr, 2007; Furnas, Landauer, Gomez, & Dumais, 1987). Like Bonita and Belle, they may differ in perspective (Barsalou & Sewell, 1984; Keysar, Lin, & Barr, 2003), intentions (Barresi & Moore, 1996), knowledge (Chi *et al.*, 1981; Gauthier *et al.*, 1998; Tanaka & Taylor, 1991; Murphy & Wright, 1984; Schvaneveldt *et al.*, 1985), and goals (Barsalou, 1983; Ratneshwar *et al.*, 2001). Thus, each actor might approach a novel activity with differing expectations (*c.f.*, Murphy & Medin, 1985) and use differing heuristics (*cf.* Lin & Murphy, 1997). For

example, Bonita uses the heuristic *big => wide*, while Belle uses the heuristic *big => long*. With each iteration of proposal/counter-proposal and ratification/clarification, each actor relies increasingly on publicly available information, shifting from private to public categories (*cf.*, Fussell & Krauss, 1989; Krauss & Fussell, 1991; Clark & Brennan, 1991).

### ***Joint Attention, Joint Intention, and Joint Reference***

With each iteration of proposal/counter-proposal and ratification/clarification, actors establish joint attention to publicly relevant features of the referent (Clark, 1972; Clark *et al.*, 1983; Tomasello & Farrar, 1986; Kaplan & Hafner, 2006; Tomasello, 1999). In particular, actors direct each other's attention (Kronmüller & Barr, 2007; Metzing & Brennan, 2003; Baldwin, 1991) to those features that allow each of them to differentiate the target referent from other possible referents (E.V. Clark, 1987; Mervis & Crisafi, 1982; Murphy & Brownell, 1985) and to infer the referent's significance to the activity (Brown, 1958; Gluck & Corter, 1985). For instance, Bonita directs Belle's attention to the widening curve of the chef's knife, which both differentiates it from the carving knife and, perhaps, suggests its function as a precise cutting tool. Further, actors establish the joint intention to act on the referent to accomplish mutual goals (Clark & Lucy, 1975; Francik & Clark, 1985; Tomasello *et al.*, 2005). Bonita and Belle jointly intend the sharpening of the chef's knife. More importantly, having established joint reference to "the half-bow knife," Bonita and Belle are likely to reuse this precedent (Brennan & Clark, 1996), reducing the effort of sharing attention and intentions on subsequent conversational turns (Clark & Wilkes-Gibbs, 1986).

### *From Joint Reference to Referential Convention*

Often, over the course of a conversation, these referential precedents develop into mutually-known, mutually-salient, and mutually-expected associations between reference and concept—*i.e.* a referential convention. Such *ad hoc* conventions derive from opportunistic and/or pre-conscious learning coupled with the dynamics of conversation. In the scenario, Bonita ratifies Belle’s proposed conceptualization “it’s like half a bow” with the name-like reference “the half-bow knife.” This might represent opportunistic learning, where Bonita exploits the *opportunity* for repeated success in referring to the chef’s knife. This might also represent pre-conscious learning, where Bonita speaks what she hears due either to the common coding of comprehension and production or some other priming mechanism (see Pickering & Garrod, 2004, 2006 on pre-conscious linguistic coordination; also see Prinz, 1990, and Liberman & Whalen, 2000 on common coding). In either case, a recently used reference is frequently reused, and a frequently used reference is mutually available and mutually expected for further reuse (Brennan & Clark, 1996). This process drastically reduces the lexical variability between communicating actors, often leading to referential conventions that one could not predict from common usage patterns and/or normative theories (Barr & Keysar, 2002; Brennan & Clark, 1996; W. Deutsch & Pechmann, 1982; Pechmann, 1989).

One common example is the over-specification of reference—*i.e.* using a reference that is more specific than the situation requires. Grice’s *Maxims of Quantity* exhort the speaker to “1. [m]ake your contribution as informative as is required (for the current purposes of the exchange) [, and] 2. [d]o not make your contribution more informative than is required” (1975, p. 45). Nevertheless, having established a

subordinate-level precedent—*e.g.*, *tagine*—communicating actors expect each other to adhere to that precedent even when a basic-level reference—*e.g.*, *brazier* or *pot*—would convey information sufficient for the joint activity (Brennan & Clark, 1996; Barr & Keysar, 2002). Moreover, these over-specified references usually persist into subsequent conversations with different conversational partners (*ibid.*).

### ***From Public to Private Categories***

A phenomenon like persistent over-specification seems unremarkable if one accepts that reference does more than label or index objects, actions, and events. In fact, a “public” category can persist beyond the confines of a conversation. Communicating actors often continue to use the “public” categories even in extra-linguistic activities that each performs in private. Studies by Markman & Makin (1998) and by Malt & Sloman (2004) illustrate this phenomenon more forcefully than would another imaginary scenario with Alpha and Beta or Bonita and Belle talking to each other.

Markman & Makin examined the effects of referential communication on category coherence beyond what a person learns from individual activity or perceptual similarity. They divided participants into three groups. In the *sort-only* group, individual participants simply sorted Lego blocks. In the *build-sort* group, individuals first built a Lego model, then sorted the blocks. Finally, in the *joint-build* group dyads collaborated in building a Lego Model, then sorted on their own. Before the building task, each dyad had to negotiate a referential lexicon for the various Lego blocks. During the building task, one member of each dyad built the Lego model while the other member described the pictorial directions (which the builder could not see).

Dyads generally adhered to their negotiated referential conventions. Also, while each dyadic lexicon differed in its particulars, they shared similar mappings between lexical references and the structural hierarchy of Lego blocks. The lexemes tended to refer to basic-level categories of blocks, and modified lexemes tended to refer to subordinate categories of those blocks. More importantly, “public” categorizations appear to percolate into private conceptualizations.

Absent the model-building context or any conversation, the *sort-only* group sorted blocks differently from one another. The experience of building the Lego model moderated variability in the sorts produced by the *build-sort* group. Adding referential communication to the building process reduced both within-dyad and between-dyad variability most of all; participants in the *joint-build* group sorted blocks much like their partners and much like members of other dyads.

Malt & Sloman (2004) observed similar effects when they examined the propagation of referential conventions through a serially conversing activity group. Participants jointly arranged photographs of common artifacts, each identifiable by two equally common or “balanced” names—*e.g.* bucket *vs.* pail and trashcan *vs.* wastebasket (see Malt & Sloman, 2004, for how they determined “balance”). Participants collaborated once with a confederate and once with each other. The confederate served as the first speaker, introducing one of the two balanced names—*e.g.* “bucket” instead of “pail”—for each of the target artifacts to the first-round participant. The first-round participant then served as speaker in collaboration with a second-round participant. The second-round participant then served as speaker in collaboration with the confederate. As elsewhere (*e.g.* Garrod & Deherty, 1994), when participants switched roles from addressee to



speaker, they usually attempted to transfer the referential conventions established in conversation with the previous speaker to conversations with subsequent addressee.

Following the collaborative card sorting, participants provided individual preference, typicality, and similarity ratings for each of the name pairs and the various target objects. Malt & Sloman found that, both when engaged in the joint activity and following the joint activity, participants overwhelmingly preferred to use the names introduced by the confederate over the other equally valid and equally common artifact names. Further, participants judged activity-related artifacts as more typical of categories named in conformity with the referential convention. Finally, participants judged activity-related artifacts as more similar to imagined prototypes of categories named in conformity with the referential convention. Again, “public” categorizations appear to percolate into “private” conceptualizations.

### **Conclusion and Hypotheses**

This overview of the literature on “private” category learning and on referential communication provides an integrated perspective on “public” categorization and the emergence of shareable referential conventions that one could not derive from either research tradition alone. The categorization literature exposes several constraints—including structure, complexity, language, and activity—that limit or enhance category learning and use among isolated individuals. The communication literature elaborates the process by which “private” conceptualizations become “public” conceptualizations and vice versa. Few studies have examined whether and how individual-level constraints operate among communicating actors. The present study looks explicitly at the

interaction of referential communication and the structure of perceptual features on the joint processes of inventing a referential lexicon for novel objects and discovering the functional significance of those objects during a novel activity. Specifically, I test three hypotheses:

***H1. Referential communication enhances indirect category learning***

This hypothesis is original to the present study. Previous research (Markman & Makin 1998) comparing “public” versus “private” category learning used stimuli for which there were multiple potential category configurations. So, one could not discern whether “public” categorizations were in any way better than “private” categorizations, only whether “public” and “private categorizations were different from one another.

Nevertheless, referential communication directs the joint attention and joint actions of communicating actors. Moreover, reference can be used to cue and compress concepts. Thus one should expect those engaged in a “public” categorization activity to learn unlabeled categories better than those engaged in a “private” categorization activity.

***H2. Referential communication enhances the learning of simple rules more than complex rules***

This hypothesis is also original to the present study. Again, previous research (Markman & Makin 1998) comparing “public” versus “private” category learning did not further compare the learning of simple versus complex rules. Moreover, isolated individuals engaged in indirect category learning tend to use both simple and complex rules with comparable facility (Minda & Ross, 2004).

Nevertheless, communicating actors may articulate and comprehend simple relationships between the referent's features and its significance to the activity more quickly and easily than complex relationships. Moreover, they may more quickly and easily cue and compress those concepts. Thus, one should expect those engaged in a “public” categorization activity to learn simple relationships more quickly and easily than complex relationships.

### ***H3. Referential communication generates conceptual homogeneity***

Communicating actors tend to conceive of activity related objects in ways similar to their partners (Garrod & Anderson, 1987; Brennan & Clark, 1996; *et al.*) and in ways similar to others engaged in the same joint activity (Markman & Makin, 1998). The present study should replicate this phenomenon. In other words, one should expect greater conceptual homogeneity among those engaged in a “public” categorization activity than among those engaged in a “private” categorization activity.

### III. METHOD

#### Purpose & Overview

With the present study, I investigated the interaction of referential communication and the structure of perceptual features on the joint processes of inventing a referential lexicon for novel objects and discovering the functional significance of those objects during an indirect category learning activity. To that end, participants worked either individually or as cooperative<sup>2</sup> dyads to learn four combinations of orthogonal functional features—nutritive *vs.* not nutritive and destructive *vs.* not destructive—that defined four categories of fictional extra-terrestrial creatures. These categories were not specifically identified or labeled; rather, participants had to learn them *indirectly* as they predicted the functions. Also, these functionally defined categories possessed a complex perceptual structure: a unidimensional (simple) rule predicted one function, while a family resemblance (complex) sub-structure predicted the other function. This function-learning task yielded function prediction data. In addition to the main function-learning task, each learner worked individually to sort the creatures (pre- and post-function learning) and to predict their functions in an individual function prediction posttest that also yielded selective attention data.

---

<sup>2</sup> Dyads had positively interdependent goals: the success of each actor depended on the other (Deutsch, 1949; Johnson & Johnson, 1989).

Together with data extracted from the transcripts of the dyadic conversations, the prediction data, sort data, and selective attention data demonstrate the extent to which:

*H1.* referential communication enhances indirect category learning;

*H2.* the learning advantages of dyadic learners vary with the complexity of the category structure; and

*H3.* referential communication generates conceptual homogeneity.

### **Participants**

Thirty-five male and thirty-seven female students (mean age 25.1) from throughout the Columbia University community participated in this study for a cash payment.

Participants were recruited using flyers posted widely across the university campus. The flyer promised a cash payment for participation and a digital audio player for the best performing participant. All participants were native speakers of the English language, with an average of four years of post-secondary schooling during which they devoted an average of two (or fewer) hours per week to computer games.

### **Design**

Three independent variables were manipulated in the design of this study.

*Learning context* served as a between-subjects factor with three levels—*dyadic learning* (N=32) versus *individual learning with 160 trials* (N=16) versus *individual learning with 320 trials* (N=24). The two individual learning conditions were created in response to difficulties in establishing an equivalent control condition to dyadic learning. Each dyadic

learner described stimuli on 160 learning trials and heard descriptions of stimuli on an additional 160 learning trials. Moreover, they received corrective feedback on predictions for all 320 trials. One could debate whether acting on a verbal description of an unseen stimulus leads to the same learning as both seeing and acting on a stimulus. So, the two individual learning conditions each represent the two boundary conditions: 160 versus 320 learning trials. Learning trials were divided into blocks of thirty-two trials. *Block* served as a within-subjects factor with either five or ten levels, depending on which learning conditions were compared. Finally, *type of rule* served as a within-subjects factor with two levels—*simple rule* versus *family resemblance*.

Log files generated during each task for each participant provided data on the effects these manipulations had on four dependent variables that capture aspects of category learning:

1. the *accuracy* with which participants predicted the functional features;
2. the *structural similarity* of participant sorts to the "true" category structure, its sub-structures, or the sorts of other participants;
3. the *type of explanation* participants cited for their sorts, and
4. the *attention allocated* to surface features when making predictions.

In addition, transcripts of the dyadic conversations provided data on *the probability with which participants referred to the surface features*.

## Materials

### *The Design of the Category Structure*

Six observable dimensions (O1-O6) and two functional dimensions (F1 and F2) defined the category structure used in this study (Table 1). The four combinations of binary values on F1 and F2 defined four categories, while sixteen combinations of binary values on O1-O5 defined the sixteen exemplars. The distribution of exemplar-defining value combinations met several criteria: each category entailed four unique exemplars, among which similar value combinations on O1-O3 predicted the value on F1 (*family resemblance* rule type), the value on O4 predicted the value on F2 (*simple* rule type), while the value on the O5 did not predict the value on any other dimension. Adding a modicum of naturalistic noise, O6 varied randomly across a range of possible values. The values on O6 did not predict the value on any other dimension.

Table 1. Category Structure represented in binary notation.

	Observable Dimensions						Functional Dimensions	
	Diagnostic				Non-Diagnostic			
EXEMPLAR	O1	O2	O3	O4	O5	O6	F1	F2
ND.1	1	1	1	1	1	random	1	1
ND.2	1	1	0	1	1	random	1	1
ND.3	1	0	1	1	0	random	1	1
ND.4	0	1	1	1	0	random	1	1
Nx.1	1	1	1	0	0	random	1	0
Nx.2	1	1	0	0	0	random	1	0
Nx.3	1	0	1	0	1	random	1	0
Nx.4	0	1	1	0	1	random	1	0
xD.1	1	0	0	1	1	random	0	1
xD.2	0	1	0	1	1	random	0	1
xD.3	0	0	1	1	0	random	0	1
xD.4	0	0	0	1	0	random	0	1
xx.1	1	0	0	0	0	random	0	0
xx.2	0	1	0	0	0	random	0	0
xx.3	0	0	1	0	1	random	0	0
xx.4	0	0	0	0	1	random	0	0

Observable dimensions O1-O3 jointly predict functional dimension F1. Observable dimension O4 predicts functional dimension F2. The assignment of the perceptual features of stimuli to O1-O4 and the functional features to F1-F2 was counterbalanced.

### *The Design of the Stimuli*

Computer graphics of fictional extra-terrestrial creatures instantiated this category structure (using the Squeak variant of Smalltalk, in a 32-bit graphical environment). Each creature possessed four diagnostic observable features—tentacles (T), fins (F), torso animation (A), and eyes (E)—which varied between two levels (Figure B1 illustrates the two levels for each physical feature). Four permutations of these physical features—(1) TFAE, (2) ETFA, (3) AETF, (4) FAET—were assigned to dimensions O1-O4 in order to counterbalance any effects due to prior expectations of physical/functional correlations. The fill colors and border colors (dimension O5) assigned to these physical features remained either stable (one of two non-predictive values) or constant. Also, the size (dimension O6) of each rendered creature varied in scale, from 90 to 110 percent (in increments of 1%) of the basic creature size. This subtle source of variation was meant to add a modicum of naturalistic noise to the category structure.

Participants encountered these creatures in the context of a computer-based video game (installed on Apple iMac computers). The scenario (Figure C1) for the game presented the learning task and the pre- and post-tests as training materials for an imaginary mission to another planet, where the “astronaut” would encounter extra-terrestrial creatures. Creatures might or might not offer *nutritive* value (described as a “jelly” that could serve either as food or bio-fuel) and might or might not be *destructive* (described as damage to “life support systems”). Successful game play required



participants to learn how to discriminate between four functionally-defined creature categories—nutritive & destructive (ND), nutritive only (NX), destructive only (XD), or no functional significance (XX). The counterbalanced assignment of functional features to dimensions F1 and F2 was meant to account for any effects due to function/rule-type pairings.

## Tasks

### *The Function Prediction Task*

Across learning contexts, participants used the Function Prediction task to learn the functional combinations that defined the four unlabeled categories of creatures (see Minda & Ross, 2004, and Monos, 1997, for similar tasks). Participants could learn the four categories indirectly by learning the functional combinations. A time-stamped log of their predictions provided data on both their *prediction accuracy*. The Function Prediction task entailed two separable roles: the *spotter*, who saw the creature; and the *beamer*, who performed the prediction-related action. In the *dyadic learning context*, each participant's initial role was assigned randomly on the first Function Prediction trial and alternated on each subsequent trial. In each of the *individual learning contexts*, each participant played both roles.

On each trial, one of the sixteen creatures appeared at the center of the spotter's Function Prediction interface (Figure C2 A & B), where it remained until an action was executed or twenty seconds had elapsed. If playing with a partner, the spotter had fifteen seconds to describe the creature to the beamer, leaving the beamer five seconds (each

second marked by a tone) to decide which functional combination—ND, NX, XD, or XX—the creature entailed and take the functionally-appropriate action—respectively, *stun & capture*, *stun*, *capture*, or *pass*. Across learning contexts, the beamer executed these actions with keystrokes on a standard computer keyboard. Individually, the <CTRL> key meant “activate tractor beam,” the <OPTION> key meant “activate stun beam,” and the <SPACE> key meant “fire” (execute the selected action). Thus, <CTRL-OPTION-SPACE> executed *stun & capture*, <OPTION-SPACE> executed *stun*, <CTRL-SPACE> executed *capture*, and <SPACE> executed *pass*. An instructions screen offered participants detailed descriptions of key combinations (Figure C3 A & B); a color-coded summary of key combinations (Figure C2 A & Figure C4) appeared along the bottom of the beamer’s Function Prediction interface.

In response to each key combination, the background of the Function Prediction interface flashed the action-appropriate color (as defined by the color-coded summary), after which visual and aural feedback signaled the functionally-related consequences of the chosen action. First, either a *positive* or *negative* tone indicated whether or not the participant executed a functionally appropriate action. Then, a synthesized voice (specifically, the IBM Expressive Speech System, Hamza, Bakis, Eide, Picheny, & Pitrelli, 2004) described the functionally-related consequences of the chosen action. For example, after capturing a NX creature, participants heard “jelly extracted;” alternatively, participants who stunned & captured an NX creature, heard “stun beam wasted, some jelly extracted.” Descriptive feedback reinforced even partially correct predictions, while correcting mistaken predictions. Table A specifies the descriptive feedback for each function by prediction pairing. Finally, a graphical energy meter (*e.g.*, Figure C4)

provided additional positive or negative feedback, namely a five to fifteen unit (pixel) increase or decrease in energy (length). As with the descriptive feedback, the energy meter reinforced even partially correct predictions, either granting partial points or dampening penalties. Table A specifies the rewards and penalties for each function/prediction pairing. The Function Prediction task continued for either five or ten blocks of trials (depending on learning context), during which each of the sixteen creatures appeared twice in random order.

### *The Sorting Tasks*

An individually-performed sorting task preceded (PRE) and followed (POST) the Function Prediction task. The two sort tasks differed only in what knowledge about functional features was available to participants. During the PRE-sort task, participants had no knowledge about the functional significance of the creatures; during the POST-sort task, participants could use what knowledge they had gleaned from the Function Prediction task. At the start of each sort task, all sixteen creatures were rendered at 15% of their normal size (in order to fit on the computer screen; see Figure C5) and were presented randomly in a two-column graphical container (along the left hand side of the screen). While one could discern observable features at the reduced size, dragging a creature onto the desktop allowed participants to view it at full size. Participants were encouraged (see sorting instructions, Figure C6 A & B) to perform a full-size inspection before deciding on how to sort each creature. In order to sort creatures, participants created a number of graphical containers into which they dragged the creatures they believed belonged together. Upon dragging a creature into a newly-created sorting

container, an explanation field appeared in the container, into which participants were instructed to provide a short explanation of why the creatures in the container belonged together. Participants could explain or edit that explanation at any time during the sorting process. Participants were free to create as many or as few categories (from one to sixteen) as they deemed necessary.

### ***The Attention Allocation Task***

The Attention Allocation task captured data on the attention participants allocated to the various physical features when deciding which functional combination—ND, NX, XD, or XX—the creatures entailed. The Attention Allocation task replicated the single-player Function Prediction task in all aspects except that each creature appeared with its various physical features hidden by graphical blinds (Figure C7). In order to uncover a feature, participants mouse-clicked its blind. Participants were instructed to uncover as many parts as they needed in order to decide each creature’s functional significance—ND, NX, XD, or XX—and whether to stun & capture, stun, capture, or pass (for instructions, screen see Figure C8). Also unlike the Function Prediction task, the Attention Allocation task lasted for a single block of thirty-two trials, during which each of the sixteen creatures appeared twice in random order.

### **Procedure**

Participants were scheduled as pairs for “gaming” sessions, each of which had been randomly designated as *dyadic*, *individual-160*, or *individual-320* sessions and

assigned one of the four physical permutations and one of the two functional permutations. After granting informed consent, participants completed a short questionnaire, providing their age, gender, years of post-secondary schooling, and the number of hours per week devoted to computer-game play. They were then seated at a computer terminal on either side of a 5'x5' barrier, beyond which each participant could hear but not see the other. There, each observed a demonstration of the “game” interface and was given time to practice the use of the keyboard and mouse. Participants then proceeded through each of the four tasks: PRE-sort, function prediction, POST-sort, and attention allocation. Before each task, participants were permitted to seek clarification of the on-screen instructions. After completion of all tasks, each participant was debriefed.

## **Data Analysis**

### ***Deriving measures of prediction accuracy***

For each individual learner, the Function Prediction task logs recorded the stimulus (represented as a vector of values on dimensions O1-O4 and F1-F2) that was presented on each trial along with the functional combination—*nutritive & destructive*, *nutritive only*, *destructive only*, or *no functional significance*—with which the learner responded to the stimulus. Similarly, for each dyad, the Function Prediction task logs recorded each stimulus that the learner “spotted” or heard described along with the response provided by the learner or his or her partner. As with the Function Prediction logs for individual learners, the Attention Allocation task logs for both dyadic and individual learners recorded the stimuli and responses of each participant.

Averaging correct responses (correctly predicted functional combinations) across blocks of thirty-two trials (two presentations of each stimulus) yielded the *functional-category prediction accuracy rates* for individuals, dyads<sup>3</sup>, and, during the Attention Allocation task, for individual dyadic learners. The same procedure for correct predictions of each function yielded *function-prediction accuracy rates*.

### ***Deriving measures of attention allocation***

In addition to the prediction data, the Attention Allocation task logs recorded what features each learner uncovered before responding with a function prediction. For each learner, these data yielded the average number of features uncovered per stimulus. Further, these data were converted into two probabilities: the probability of uncovering a *family resemblance* feature and the probability of uncovering a *simple rule* feature. Specifically, the number of *family resemblance* features uncovered by each learner was summed across Attention Allocation trials then normalized by the maximum number of *family resemblance* features the learner could have uncovered. The probability of uncovering a *simple rule* feature was derived in the same way.

---

<sup>3</sup> The responses of dyad members are dependent on one another during the Function Prediction task; thus, the dyad rather than the dyad member serves as the unit of analysis.

***Deriving types of sorting explanations & measures of structural similarity***

The logs recorded during the sorting tasks that preceded (PRE) and followed (POST) the Function Prediction task provided data on which creatures were sorted together as well as the explanations participants used to justify these creature clusters.

The explanations were segmented and coded as either function/behaviorally related (citing  $\pm$  nutritive and  $\pm$  destructive and/or  $\pm$  capture and  $\pm$  stun) or perceptually related (citing the surface features). Additionally, both function-related and perceptually-related explanations were coded as either *family resemblance* related (citing a feature or function related to the *family resemblance* substructure) or *simple rule* related (citing a feature or function related to the *simple rule* substructure). Normalizing the frequency of each type of explanation by the number of creature clusters times the number of explanation types yielded the probability of mentioning that type of feature. Similarly, normalizing the frequency of explanations that cited functional/behavioral combinations by the number of creature clusters yielded the probability of mentioning a *functional category*.

To derive measures of *structural similarity*, each participant's post-learning creature groups were converted into binary co-occurrence matrices. Three additional matrices, represented co-occurrence based on (1) the "true" category clusters, (2) the destructive-function category clusters, and (3) the nutritive-function category clusters. The lower triangle (the binary values below the diagonal) of each co-occurrence matrix was rearranged as a vector. The jaccard similarity (which is appropriate for binary presence/absence data; Jaccard, 1912) between the various co-occurrence vectors served

as an indicator of *structural similarity*. *Structural similarity* was used to test inter-participant similarities as well as the effects of *learning context* and *type of rule*.



## IV. RESULTS

Overall, the results demonstrate that dyadic learners learned the functionally defined categories more quickly and more accurately than individual learners. Also, dyadic learners exhibited greater conceptual homogeneity than individual learners. Nevertheless, the dyadic advantage in learning functions appeared earlier and was greater for the *simple-rule* predicted function than for the *family-resemblance* predicted function.

### Referential communication enhances indirect category learning

#### *Functional-category prediction accuracy*

Dyadic learners predicted functionally defined categories with greater accuracy than did individual learners. Participants in the two individual learning contexts were similarly accurate in predicting functional categories during the initial five blocks of learning trials (1-160)<sup>4</sup>. This dyadic advantage is evident in Figure 1 (below), where three accuracy curves compare functional-category prediction accuracy rates by *learning context* across the initial five *blocks* of learning trials (1-160) and across the final five *blocks* of learning trials (161-320).

Moreover, this early dyadic advantage was corroborated by a repeated-measures ANOVA on the functional-category prediction accuracy rate with *learning context* as a

---

<sup>4</sup> Unless otherwise noted, data from the two individual learning conditions were combined for statistical tests relating to learning trials 1-160.

between-subjects factor and *block* (the five blocks that entailed learning trials 1-160) as a within-subjects factor. The interaction of *learning context* and *block* was significant— $F_{(4,216)} = 10.522$ ,  $p < .000$ ,  $\hat{\eta}^2 = .163$ —as were the main effects of *learning context*— $F_{(1,54)} = 14.839$ ,  $p < .000$ ,  $\hat{\eta}^2 = .216$ —and *block*— $F_{(4,216)} = 21.021$ ,  $p < .000$ ,  $\hat{\eta}^2 = .280$ . On average across the first 160 learning trials, dyadic learners predicted functional categories with *an accuracy rate that increased more steeply* than the accuracy rate of individual learners.

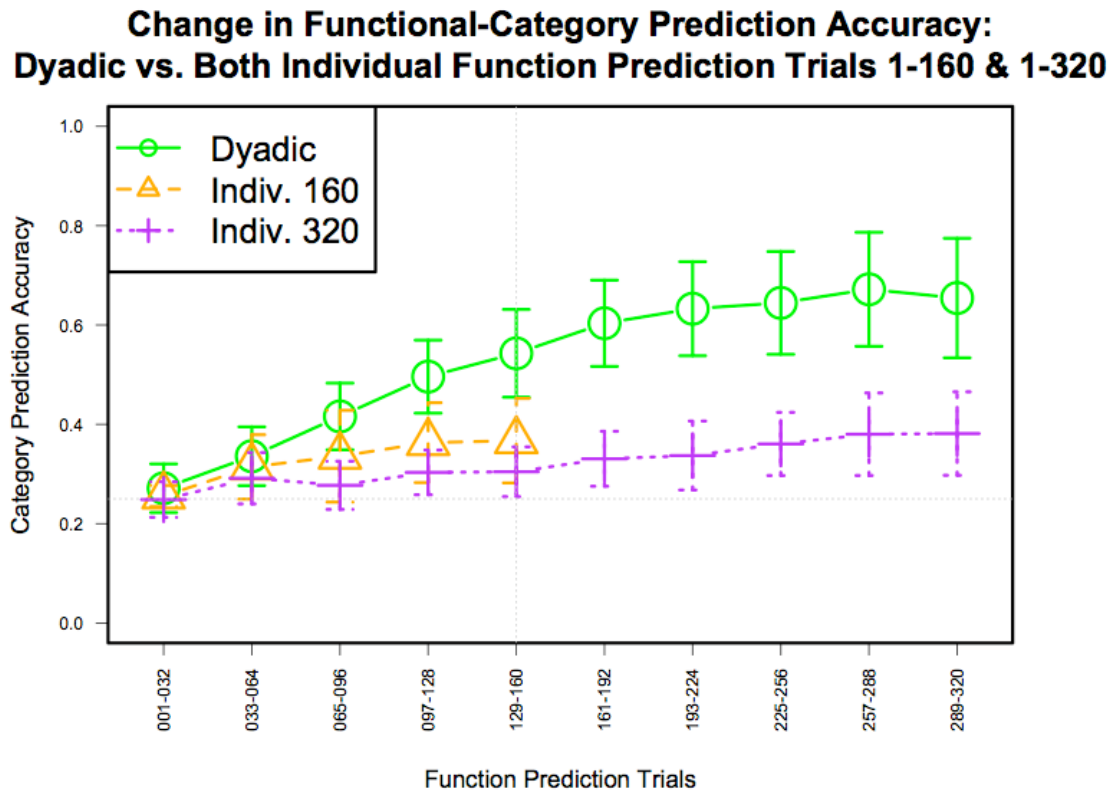


Figure 1. Compares the rate of functional-category prediction accuracy of dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials). Error bars indicate confidence interval.

The continued dyadic advantage was also corroborated by a repeated-measures ANOVA on the functional-category prediction accuracy rate with *learning context*<sup>5</sup> as a between-subjects factor and *block* (the five blocks that entailed learning trials 161-320) as a within-subjects factor. Again, the interaction of *learning context* and *block* was significant— $F_{(4,152)} = 8.980$ ,  $p < .000$ ,  $\hat{\eta}^2 = .191$ —as were the main effects of *learning context*— $F_{(1,38)} = 30.858$ ,  $p < .000$ ,  $\hat{\eta}^2 = .448$ —and *block*— $F_{(4,152)} = 23.804$ ,  $p < .000$ ,  $\hat{\eta}^2 = .328$ . On average across the final 160 learning trials, dyadic learners predicted functional categories with *an accuracy rate that continued to increase more steeply* than the accuracy rate of individual learners.

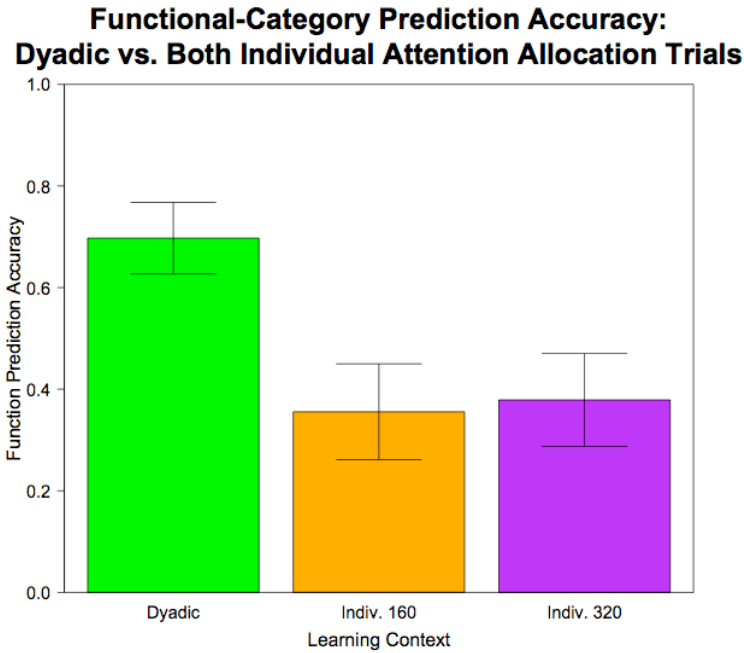


Figure 2. Compares the rate of functional-category prediction *accuracy* of dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) during the Attention Allocation task. Error bars indicate confidence interval.

---

<sup>5</sup> Unless otherwise noted statistical tests relating to learning trials 161-320 compared dyads only to individuals who completed 320 learning trials.

**Average Number of Features Uncovered Per Stimulus:  
Dyadic vs. Both Individual Attention Allocation Trials**

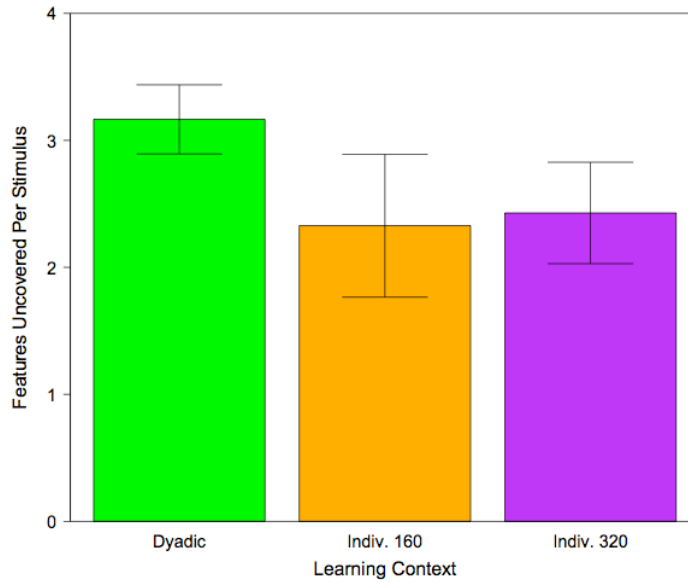


Figure 3. Compares the average number of features per stimulus that dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) uncovered during the Attention Allocation task. Error bars indicate confidence interval.

### ***Posttest category prediction accuracy and attention allocation***

The dyadic advantage persists into the attention allocation task (Figure 2). Not surprisingly, attention allocation—as represented by the average number of features uncovered per stimulus (Figure 3)—correlated with functional-category prediction accuracy— $r_{(acc.,attend)} = .471$ ,  $p < .000$ . A MANOVA on the functional-category prediction accuracy rate and attention allocation (features per stimulus) with *learning context*<sup>6</sup> as a between-subjects factor yielded a significant multivariate effect— $F_{(2,69)} = 26.669$ ,  $p < .000$ ,  $\hat{\eta}^2 = .249$ —as well as significant univariate effects on functional-category prediction accuracy— $F_{(1,70)} = 51.022$ ,  $p < .000$ ,  $\hat{\eta}^2 = .422$ —and on attention allocation—

---

<sup>6</sup> Unless otherwise noted, data from the two individual learning conditions were combined for statistical tests relating to the post-learning tasks.

$F_{(1,70)} = 14.275$ ,  $p < .000$ ,  $\hat{\eta}^2 = .169$ . On average, during the attention allocation task, dyadic learners predicted functional categories with *greater accuracy* than individual learners— $M_{(dyad)} = .697$  ( $SD = .192$ ) vs.  $M_{(indiv.)} = .370$  ( $SD = .195$ )—while uncovering a *greater number of perceptual features* than individual learners— $M_{(dyad)} = 3.165$  ( $SD = .743$ ) vs.  $M_{(indiv.)} = 2.388$  ( $SD = .954$ ).

Further, participants needed to uncover at least three features to determine the functional category of a stimulus. The mean number of features uncovered by dyadic learners, 3.165 ( $SD = .743$ ), did not differ significantly from that minimum— $t_{(31)} = 1.256$ ,  $p = .219$ . The mean number of features uncovered by individual learners, 2.388 ( $SD = .954$ ), fell short of the minimum— $t_{(39)} = -4.057$ ,  $p < .000$ .

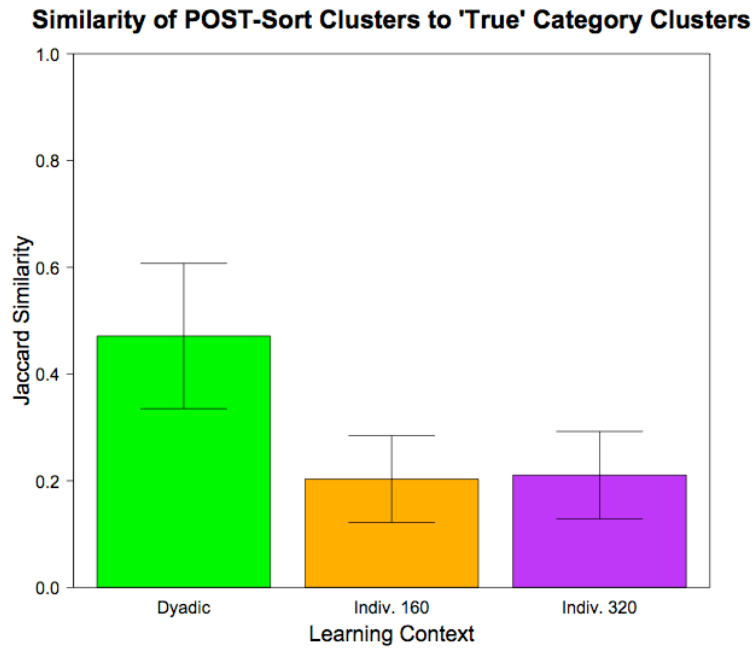


Figure 4. Compares the *structural similarity* of the post-learning sort clusters produced by dyadic learners and individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) to the “true” category clusters. Error bars indicate confidence interval.

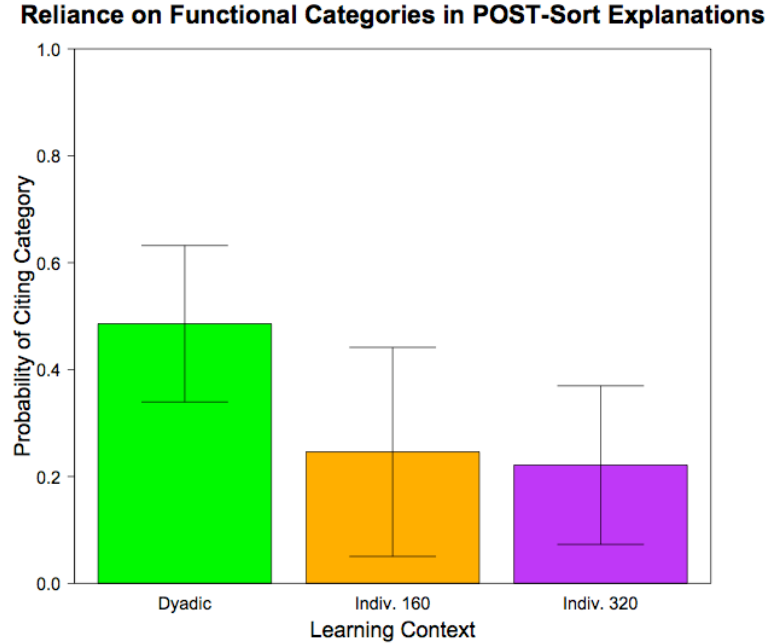


Figure 5. Compares the probability of citing a category (whether functionally-defined or behaviorally-defined) when dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) explained their post-learning sort clusters. Error bars indicate confidence interval.

### *The structure of post-learning sort clusters*

The dyadic learning advantage was also evident in how participants sorted creatures after the function prediction task. As described in the *Data Analysis* section of the *Method* chapter, post-learning sort clusters and the “true” category clusters were converted into binary co-occurrence matrices, each of which was rearranged as a vector. The jaccard similarity between each of the various co-occurrence vectors served as an indicator of *structural similarity* between clusters. As apparent in Figure 4, the post-learning sort clusters of dyadic learners were more similar to the “true” category clusters than were the post-learning sort clusters of individual learners— $M_{(dyad)} = .471$  ( $SD = .372$ ) vs.  $M_{(indiv.)} = .207$  ( $SD = .172$ ). Moreover, in explaining their sorts (Figure 5), dyadic learners cited functional and/or behavioral categories (the *beams* used in

response to stimuli) with greater likelihood than did individual learners— $M_{(dyad)} = .486$  ( $SD = .399$ ) vs.  $M_{(indiv.)} = .231$  ( $SD = .344$ ). The similarity of the post-learning sort clusters and the “true” category clusters correlated with the reliance on category-related explanations of those clusters— $r_{(sim., explain)} = .614$ ,  $p < .000$ .

The differences between learning contexts were corroborated by a MANOVA on *structural similarity* (jaccard similarity of sorted co-occurrence vectors to the “true” category co-occurrence vectors) and *type of explanation* was run with *learning context* as a between-subjects factor. The multivariate effect of *learning context* was significant— $F_{(2,69)} = 8.169$ ,  $p < .000$ ,  $\hat{\eta}^2 = .563$ —as was its univariate effect on *structural similarity*— $F_{(1,70)} = 15.881$ ,  $p < .000$ ,  $\hat{\eta}^2 = .185$ —and on *type of explanation*— $F_{(1,70)} = 8.4183$ ,  $p = .005$ ,  $\hat{\eta}^2 = .107$ . In all, dyadic learners, more than individual learners, sorted in accordance with the normative (experimenter-designed) category clusters.

### **Referential communication generates conceptual homogeneity**

#### ***Within-Dyad and Between-Dyad Sorting Homogeneity***

Dyadic learners sorted the stimuli in ways that closely resembled the “true” category clusters and, to varying degrees, the sort clusters of other learners. Figure 6 compares the pair-wise jaccard similarities between the co-occurrence vectors of dyadic learners paired with their partners ( $M = .454$ ,  $SD = .395$ ) vs. pair-wise jaccard similarities between the co-occurrence vectors of various pseudo-dyads, including: dyadic learners paired randomly ( $M = .274$ ,  $SD = .286$ ), dyadic learners paired with each

non-partner ( $M = .249$ ,  $SD = .262$ ), and individual learners paired with each other individual learner ( $M = .140$ ,  $SD = .128$ ).

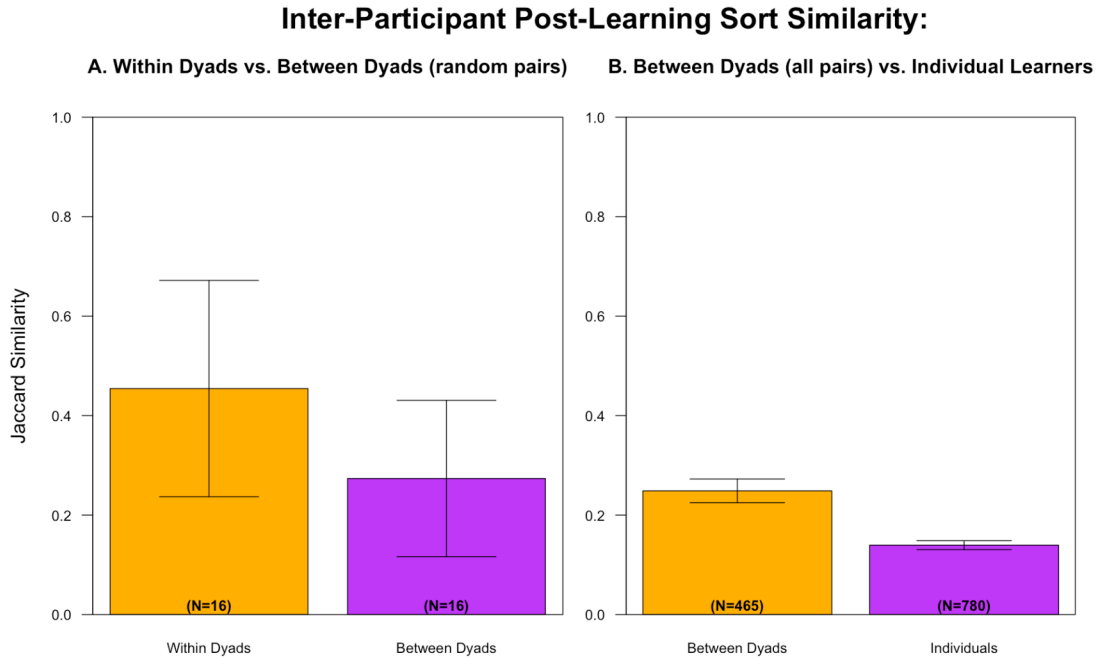


Figure 6. Compares the mean pair-wise similarity of the post-learning sort clusters produced by actual dyads vs. between-dyad pairings and pairs of individual learners. Error bars indicate confidence interval.

The post-learning sort clusters of actual dyads were not significantly more similar to one another than the sort clusters of randomly paired dyadic learners— $t_{(27.32)} = 1.483$ ,  $p = .149$ . Nevertheless, the post-learning sort clusters of non-partners were significantly more similar to one another than the sort clusters of individual learners,  $t_{(597.131)} = 8.404$ ,  $p < .000$ . One should note that, prior to learning, the sort clusters produced by dyadic learners (pair-wise jaccard similarity  $M = .192$ ,  $SD = .211$ ) were no more similar to one another than were the sort clusters produced by individual learners (pair-wise jaccard similarity  $M = .196$ ,  $SD = .171$ ),  $t_{(673.077)} = -.290$ ,  $p = .772$ .



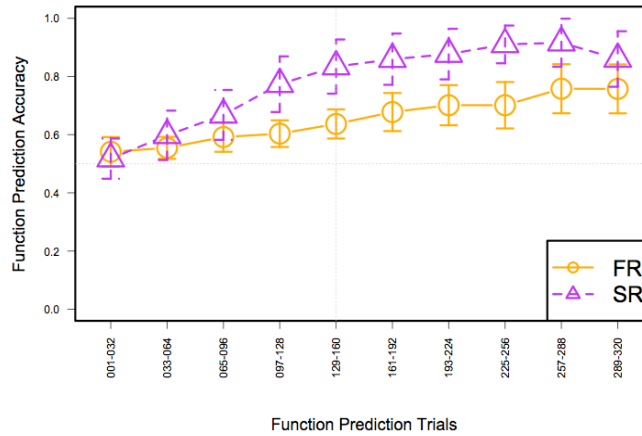
### ***Within-Dyad Referential Homogeneity***

Dyadic learners also tended to reference the same features as their partners—i.e. they used similar referential conventions (conventions were represented as a vector of probabilities with which each dyadic learner mentioned the various features). While partners mentioned the various features with differing probabilities when explaining their pre-learning sort clusters ( $median\ r_{(pre-A, pre-B)} = .422$ ) they referred to those features with highly correlated probabilities ( $median\ r_{(initial-A, initial-B)} = .994$ ) during the initial block of learning trials. Partners continued to use highly compatible conventions into the final block of learning trials ( $median\ r_{(final-A, final-B)} = .995$ ) though conventions did change slightly between the initial and final blocks of learning trials ( $median\ r_{(initial, final)} = .847$ ).

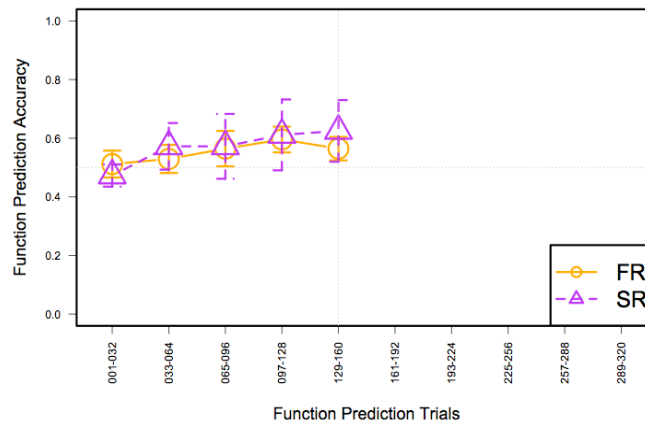
In addition, partners established reference to creatures with greater efficiency over time. Partners uttered significantly fewer words per diagnostic feature during the final block of dyadic learning trials ( $M = 2.349$ ,  $SD = .898$ ) than during the initial block of dyadic learning trials ( $M = 3.649$ ,  $SD = .925$ )— $t_{(31)} = 6.221$ ,  $p < .000$ . Finally, referential conventions accorded well with each partner's attention to the various features. The probabilities with which dyads referred to the various features during the last block of dyadic learning trials correlated strongly with the probabilities with which each partner would later uncover those features during the attention allocation task ( $median\ r_{(final, attend-A)} = .922$  and  $median\ r_{(final, attend-B)} = .965$ ). Also during the attention allocation task, partners uncovered features with highly correlated probabilities ( $median\ r_{(attend-A, attend-B)} = .981$ ).

### Change in SR-Function & FR-Function Prediction Accuracy:

#### A. Dyadic Function Prediction Trials 1-320



#### B. Individual Function Prediction Trials 1-160



#### C. Individual Function Prediction Trials 1-320

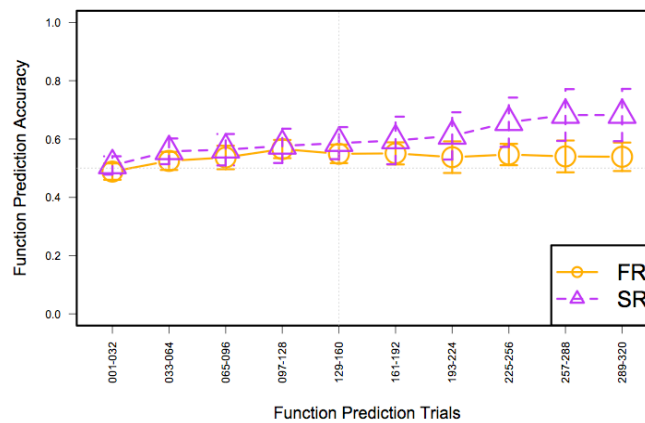


Figure 7 A-C. Compare the *simple-rule* function prediction accuracy rates vs. *family-resemblance* function prediction accuracy rates of dyadic learners vs. individual learners who completed 160 learning trials vs. individual learners who completed 320 learning trials. Error bars indicate confidence interval.

## Referential communication enhances simple-rule learning more than complex-rule learning

### *Function prediction accuracy*

On average across the first 160 learning trials, dyadic learners predicted the *simple-rule* related functions with *more sharply increasing accuracy rates* than they predicted the *family-resemblance* related functions, while participants in the two individual conditions predicted both functions with similar accuracy (Figure 7 A-C). This early prediction asymmetry for dyadic learners was corroborated by a repeated-measures ANOVA on the function prediction accuracy rate with *learning context* as a between-subjects factor and *type of rule* and *block* (the five blocks that entailed learning trials 1-160) as within-subjects factors. The three-way interaction of *learning context*, *type of rule*, and *block* was significant— $F_{(4,216)} = 3.855$ ,  $p = .011$ ,  $\hat{\eta}^2 = .067$ —as were the two-way interactions of *learning context* and *type of rule*  $F_{(1,54)} = 5.326$ ,  $p = .025$ ,  $\hat{\eta}^2 = .090$ —and *learning context* and *block*— $F_{(4,216)} = 9.218$ ,  $p < .000$ ,  $\hat{\eta}^2 = .146$ . All main effects were also significant, including: *learning context* ( $F_{(1,54)} = 14.254$ ,  $p < .000$ ,  $\hat{\eta}^2 = .209$ ), *type of rule* ( $F_{(1,54)} = 8.4169$ ,  $p = .005$ ,  $\hat{\eta}^2 = .135$ ) and *block* ( $F_{(4,216)} = 12.258$ ,  $p < .000$ ,  $\hat{\eta}^2 = .185$ ). The interaction of *type of rule* and *block* failed to reach significance— $\hat{\eta}^2 = .026$ .

Dyadic learners, more than individual learners, continued to predict the *simple-rule* related functions with *increasingly greater accuracy* than they predicted the *family-resemblance* related functions across the final 160 learning trials. A repeated-measures ANOVA on the function prediction accuracy rate with *learning context* as a between-subjects factor and *type of rule* and *block* (the five blocks that entailed learning trials 161-

320) as within-subjects factors yielded a significant three-way interaction of *learning context*, *type of rule*, and *block* was significant— $F_{(4,152)} = 2.274$ ,  $p = .029$ ,  $\hat{\eta}^2 = .056$ . The two-way interactions of *learning context* and *type of rule*— $F_{(1,38)} = 4.917$ ,  $p = .033$ ,  $\hat{\eta}^2 = .115$ —*learning context* and *block*— $F_{(4,152)} = 7.811$ ,  $p < .000$ ,  $\hat{\eta}^2 = .067$ —and *type of rule* and *block*— $F_{(4,152)} = 4.081$ ,  $p < .000$ ,  $\hat{\eta}^2 = .097$ —were also significant, as were the main effects of *learning context* ( $F_{(1,38)} = 31.328$ ,  $p < .000$ ,  $\hat{\eta}^2 = .452$ ), *type of rule* ( $F_{(1,38)} = 21.326$ ,  $p < .000$ ,  $\hat{\eta}^2 = .359$ ) and *block* ( $F_{(4,152)} = 20.635$ ,  $p < .000$ ,  $\hat{\eta}^2 = .352$ ).

Nevertheless, individuals who continued for 320 trials also exhibited a prediction asymmetry. They predicted the *simple-rule* related functions with *increasingly greater accuracy* than they predicted the *family-resemblance* related functions. A repeated-measures ANOVA on the function-prediction accuracy rate with *type of rule* and *block* (learning trials 161-320) as within-subjects factors yielded a significant interaction of *type of rule* and *block*— $F_{(4,92)} = 4.014$ ,  $p = .005$ ,  $\hat{\eta}^2 = .149$ —as well as significant main effects of *type of rule* ( $F_{(1,23)} = 9.697$ ,  $p = .005$ ,  $\hat{\eta}^2 = .296$ ) and *block* ( $F_{(4,92)} = 2.791$ ,  $p = .031$ ,  $\hat{\eta}^2 = .108$ ).

### ***Posttest function prediction accuracy and attention allocation***

During the post-learning Attention Allocation task, participants across learning contexts predicted the *simple-rule* related functions with greater accuracy than the *family-resemblance* related functions (Figure 8). Again, attention allocation—as represented by the probability of uncovering a *simple-rule* related vs. *family-resemblance* related feature—correlated with prediction accuracy— $r_{(acc.,attend)} = .519$ ,  $p < .000$  (Figure 9).

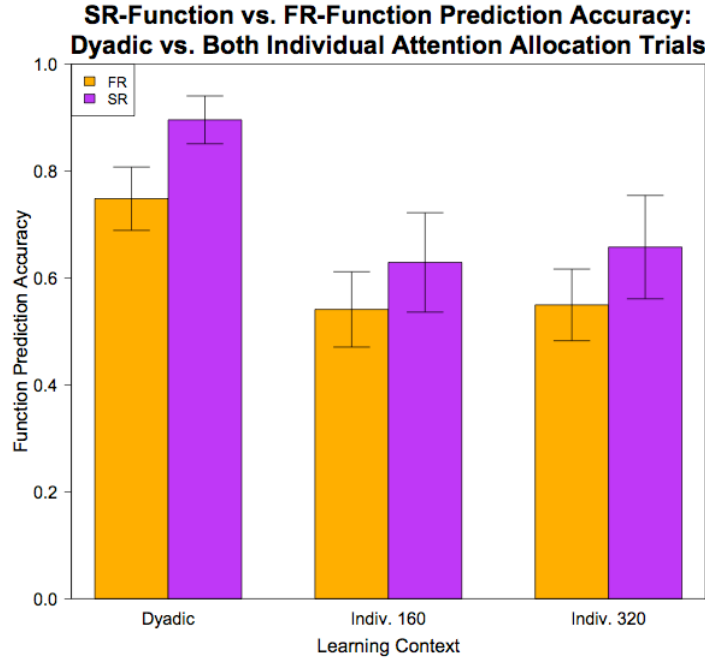


Figure 8. Compares the *simple-rule* function prediction accuracy rates vs. *simple-rule* function prediction accuracy rates of dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) during the Attention Allocation task. Error bars indicate confidence interval.

A MANOVA on the function prediction accuracy rate and attention allocation (the probability of uncovering a feature) with *learning context* as a between-subjects factor and *type of rule* as a within-subjects factor yielded a significant multivariate effect for *learning context*— $F_{(2,69)} = 26.353$ ,  $p < .000$ ,  $\hat{\eta}^2 = .247$ —and for *type of rule*— $F_{(2,69)} = 14.879$ ,  $p < .000$ ,  $\hat{\eta}^2 = .164$ , but not for the interaction of *learning context* and *type of rule*— $\hat{\eta}^2 = .012$ . On average during the Attention Allocation task, dyadic learners predicted both functions with *greater accuracy* than individual learners ( $M_{(dyad)} = .821$ ,  $SD = .160$ , vs.  $M_{(indiv.)} = .596$ ,  $SD = .181$ ), while uncovering features with *greater likelihood* than individual learners ( $M_{(dyad)} = .825$ ,  $SD = .253$ , vs.  $M_{(indiv.)} = .633$ ,  $SD = .345$ ). Nevertheless, those differences between dyadic learners and individual

learners *did not depend* on which rule—*simple* or *family resemblance*—predicted the functions— $M_{(dyad-sr)} = .896$  ( $SD = .122$ ) vs.  $M_{(dyad-fr)} = .748$  ( $SD = .162$ ) and  $M_{(indiv-sr)} = .646$  ( $SD = .201$ ) vs.  $M_{(indiv-fr)} = .546$  ( $SD = .143$ ).

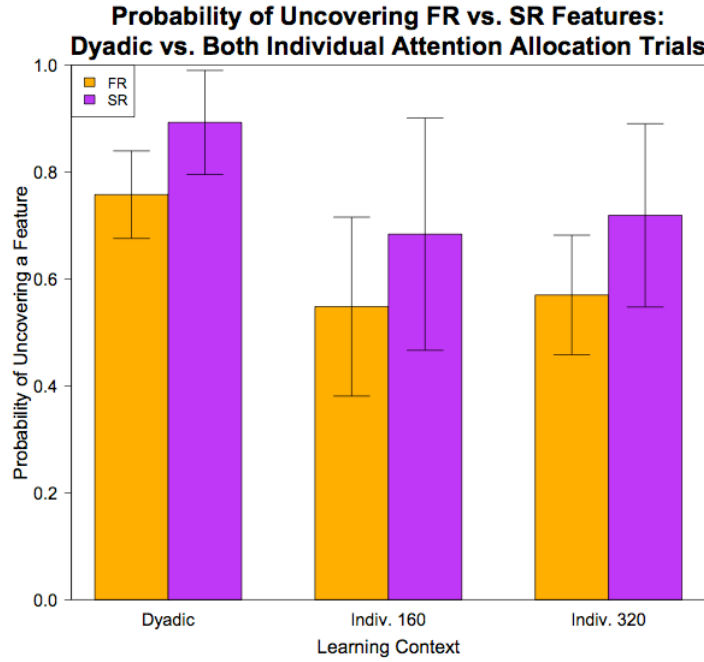


Figure 9. Compares the probability that a dyadic learner vs. individual learner (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) would uncover a *family-resemblance* related versus *simple-rule* related feature on any Attention Allocation task trial. Error bars indicate confidence interval.

Univariate effects followed the same pattern. *Learning context* had a significant main effect on the function prediction accuracy rate —  $F_{(1,70)} = 52.019$ ,  $p < .000$ ,  $\hat{\eta}^2 = .426$ —and on attention allocation— $F_{(1,70)} = 13.117$ ,  $p < .000$ ,  $\hat{\eta}^2 = .158$ . *Type of rule* had a significant main effect on the function prediction accuracy rate —  $F_{(1,70)} = 30.117$ ,  $p < .000$ ,  $\hat{\eta}^2 = .301$ —and on attention allocation— $F_{(1,70)} = 8.722$ ,  $p = .004$ ,  $\hat{\eta}^2 = .111$ . The interaction of *learning context* and *type of rule* failed to reach significance, with  $\hat{\eta}^2 = .016$  for the effect on the function prediction accuracy rate and  $\hat{\eta}^2 < .000$  for the effect on attention allocation.

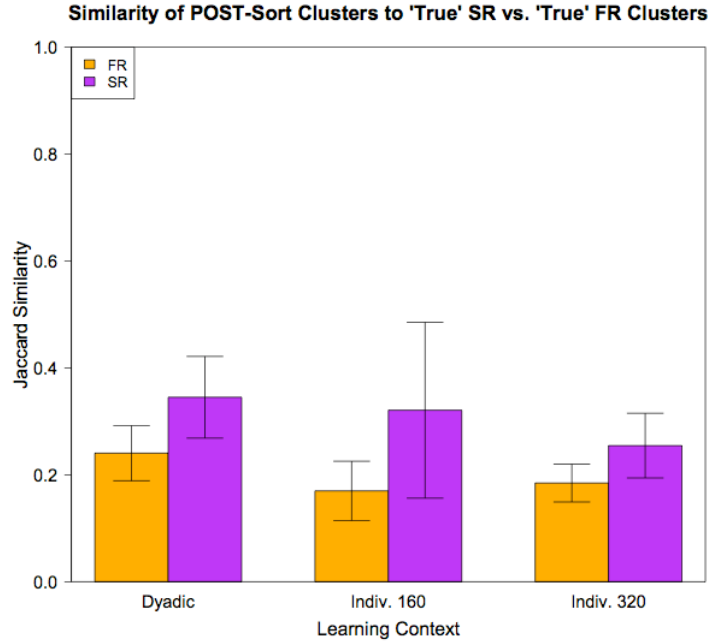


Figure 10. Compares the *structural similarity* of the post-learning sort clusters produced by dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) to the category clusters defined by one or another *type of rule*. Error bars indicate confidence interval.

### ***Structural similarity of post-learning sort clusters to the rule-defined clusters***

Similarly, participants across learning contexts sorted creatures into clusters that resembled *simple-rule* defined clusters more than *family-resemblance* defined clusters, though the sorts produced by dyadic learners were slightly more similar to the clusters defined by each type of rule (Figure 10). An ANOVA on *structural similarity* (jaccard similarity of sorted co-occurrence vectors to the *simple-rule* defined and the *family-resemblance* defined co-occurrence vectors) with *learning context* as a between-subjects factor and *type of rule* as a within-subjects factor yielded a significant main effect for *type of rule*  $F_{(1,31)} = 22.890$ ,  $p < .000$ ,  $\hat{\eta}^2 = .246$ . *Learning context* was marginally significant— $F_{(1,70)} = 3.353$ ,  $p = .071$ ,  $\hat{\eta}^2 = .046$ —and the interaction of *learning context* and *type of rule* failed to reach significance— $\hat{\eta}^2 < .000$ . Further, dyadic and individual

learners cited *simple-rule* related features and *family-resemblance* related features with comparable likelihood when explaining their post-learning sorts (Figure 11). In all, participants across learning contexts *sorted more in accordance* with the *simple-rule* defined structure than the *family-resemblance* defined structure. Nevertheless, participants across learning context did not mention *simple-rule* related features with greater likelihood than *family-resemblance* related features.

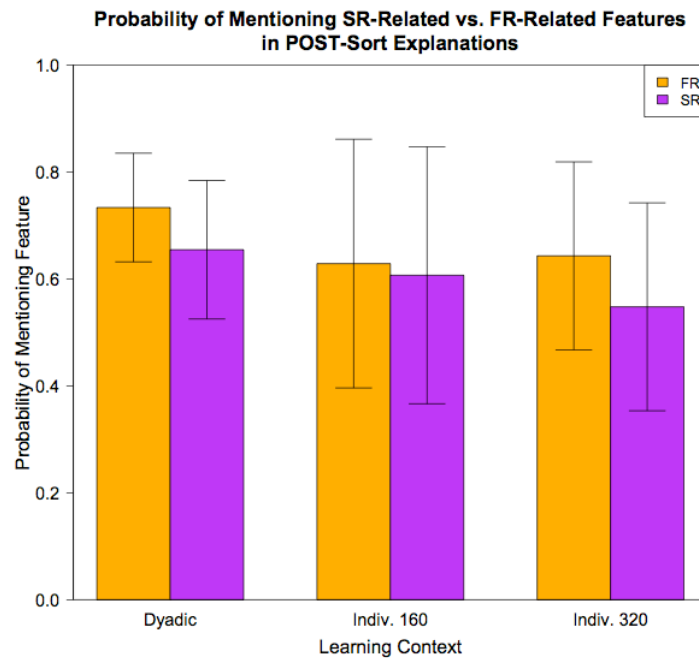


Figure 11. Compares the probability that dyadic learners vs. individual learners (both individuals who completed 160 learning trials and individuals who completed 320 learning trials) mentioned features relating to one or another type of rule when explaining their post-learning sorts. Error bars indicate confidence interval.



## V. DISCUSSION

With the present study, I investigated the interaction of referential communication and the structure of perceptual features on the joint processes of inventing a referential lexicon for novel objects and discovering the functional significance of those objects during an indirect category learning activity. As hypothesized, referential communication led to better learning of functional categories (*H1*), though earlier and more so for functions predicted by simple rather than complex rules (*H2*). These effects of communication reveal previously untested differences between “public” and “private” category learning. Moreover, referential communication reduced conceptual variability within and between collaborating dyads (*H3*), while individuals who performed the same learning tasks remained conceptually heterogeneous. This effect of communication replicates and reinforces previous research (*e.g.*, Markman & Makin, 1998).

To explain the learning advantages observed among dyadic learners, I argue that referential communication may *direct attention to relationships between features (perceptual and functional) and actions* as well as *render such relationships more memorable*. Moreover, communication may *foster and/or sustain greater motivation* among collaborators and may allow them to *take advantage of the differing expectations and heuristics* each collaborator brings to the task.

To explain the simplicity advantages observed among dyadic learners, I argue that referential communication may *provide explicit “rules” for otherwise implicit (and perhaps more difficult) judgements*. Dyads appear to have established reference to simple

rules earlier than they established reference to complex rules; thus, they could explicitly (and perhaps more easily) learn the simple rule earlier than the complex rule.

Finally, to explain the conceptual homogeneity between and within dyads, I consider whether communication *pushes “public” conceptualizations and publicly-formed “private” conceptualizations towards a limited range of widely shareable conceptual structures.*

In what follows, I summarize the results that support each of these hypotheses and elaborate the preceding explanations of the results. Further, I speculate what these results might imply for research on category learning, communication, and joint activity. Finally, I suggest ways in which future research might further explore the relationship between “public” and “private” concepts.

### **Referential communication enhances indirect category learning**

#### ***Summary of Results***

The results demonstrate that dyadic learners learned the functionally defined categories more quickly and more accurately than individual learners. Dyadic learners predicted functions with increasing accuracy across the Function Prediction (learning) trials. Only those individual learners who continued beyond 160 learning trials exhibited category prediction accuracy rates above thirty percent (where chance = 25%) as they approached the end of the learning task. Moreover, dyadic learners, more than individual learners, sorted stimuli into clusters that resembled the normative or “true” functional category clusters and relied more heavily than individual learners on functionally-defined or behaviorally-defined (*i.e.*, what combination of “beams” were required in response to

the clustered stimuli) categories in explaining their sort clusters. Finally, dyadic learners attended to (*i.e.*, uncovered) a greater number of diagnostic features than individual learners during the Attention Allocation posttest trials. In fact, dyadic learners uncovered at least as many features (a minimum of three) as required to determine the functional categories of creatures, while individuals fell short of the minimum.

### ***Explanation of the dyadic advantage***

Referential communication directs the joint attention and the joint actions of communicating dyads (*e.g.*, Kronmüller & Barr, 2007; Clark & Lucy, 1975). For example, one dyad (D10.1) used “juicy-eyed” to reference the relationship between type of eyes and the nutritive “jelly” that the beamer should “juice” or extract from the creatures. By drawing attention to relationships between perceptual and functional features (eyes => jelly) and between features and actions (eyes => jelly => capture), referring expressions may cue the concept defined by these relationships (*c.f.*, Lupyan *et al.*, 2007); in this case, a nutritive creature that the beamer must capture. Moreover, by encapsulating the relationships that define the concept, referring expressions may render such relationships more memorable (*ibid.*).

In addition to cuing a concept and the relationships that define the concept, referring expressions may serve as a compressed form of the concepts and its features—*i.e.*, a conceptual manipulative (A. Clark & Karmiloff-Smith, 1993). The referring expression “juicy-eyed” compresses the relationship between eye type, nutritive value, and the extraction of that value. By using the referring expression “juicy-eyed,” the members of D10.1 could easily infer the appropriate action, while avoiding burdensome comparisons of the different eye types and the functions and actions each type implied.

While the referring expression “juicy-eyed” is especially evocative, communicating dyads might reap similar benefits from more mundane but often repeated referring expressions. Interlocutors tend to reuse each other’s referring expressions (*e.g.*, Garrod & Anderson, 1987). Moreover, a recently used referring expression is frequently reused, and a frequently used referring expression is mutually available and mutually expected for further reuse (Brennan & Clark, 1996). As observed in the present study, these repeated referring expressions serve as *ad hoc*, yet stable, conventions of reference, with each partner of a communicating dyad referring to the same features. Such *ad hoc* conventions reduce cognitive load (Clark & Wilkes-Gibbs, 1986); the use of conventional references is more likely to convey mutually salient, mutually expected, and mutually understood conceptual content (*c.f.*, Jolicoeur *et al.*, 1984).

Taken together, these various cognitive benefits deriving from referential communication and referential convention might explain why those engaged in a “public” categorization activity learn novel categories better than those engaged in a “private” categorization activity.

That said, other aspects of social and collaborative activity might also explain the dyadic advantage observed in the present study. The learning advantages among dyadic learners may result from increased motivation. This goes beyond mere social facilitation, otherwise the prize offered to the best performing participant, audience effects (via the experimenter), and co-action effects (performing the task at the same time as other participants) should have helped individual learners, as well (*c.f.*, Zajonc, 1965). Instead, motivation may derive from the interdependence of dyadic learners (*c.f.*, Deutsch, 1949; Johnson & Johnson, 1989). That is, the “spotter’s” *score* depended on the “beamer”

making an accurate prediction, which, in turn, depended on establishing joint reference to creatures. Hence, dyadic learners were motivated to jettison idiosyncratic conceptions, expend greater effort on the task, and avoid satisficing solutions.

The dyadic advantage observed in the present study may also result from knowledge diversity or, more accurately the differing expectations and heuristics that derive from differing knowledge. Dyadic learners may approach the Function Prediction task with differing expectations and use different heuristics to yield differing and perhaps more numerous hypotheses about the relationships between perceptual and functional features and between those features and their actions. Competing hypotheses may yield better hypotheses about novel objects in a novel activity (*c.f.*, Wiley & Jolly, 2003). Diversity among partners may also encourage partners to proffer simpler and more shareable hypotheses. Interlocutors try to minimize the joint effort of sharing beliefs (Clark & Wilkes-Gibbs, 1986) and tend to take each other's knowledge into account when trying to establish joint reference (Fussell & Krauss, 1989).

Similarly, human beings may differ in their previously developed inductive reasoning skills. Such skills might provide a participant an early advantage in learning to predict functional categories. While participants were randomly assigned to learning conditions, one cannot entirely ignore the possibility that a greater number of highly skilled participants were assigned to the dyadic condition. In the present study, partners predicted functional categories with highly correlated accuracy rates across the 320 learning trials ( $median\ r_{(acc.-A, acc.-B)} = .744$ ). This similarity in prediction accuracy would suggest that either all partners were comparably skilled at induction or that the more skilled partners were also adept at circumventing the referential strictures of the task (*i.e.*,

no explicit guidance on predictions) to convey predictive information to their less skilled partners. In either case, a pretest of inductive reasoning may help future studies prevent the remote possibility of assigning a greater number of highly skilled participants to the dyadic condition.

Then again, these alternative explanations do not diminish the influence of referential communication. Only through communication could dyadic learners take advantage of interdependence and/or diversity of knowledge and/or skills. Without communication motivation may dwindle and diversity may hinder rather than help the learning process.

Differences in the tasks performed by dyadic and individual learners may also lead to a learning advantage in the dyadic condition. Individual learners perform both the role of the beamer and the role of the spotter on each learning trial; dyadic learners alternate these roles with each trial. On the one hand, one might argue that separating the roles reduces cognitive load for dyadic learners: the spotter handles the perceptual discrimination task, while the beamer handles the prediction task. On the other hand, one might also argue that each separate role in the dyadic learning condition requires cognitive effort comparable to that of the joint role in the individual learning condition. In addition to the perceptual discrimination task, the spotter must *produce* a referring expression that conveys the *predictive* perceptual information to the beamer. The beamer must *comprehend* the referring expression and *interpret* how the verbally conveyed perceptual information relates to the prediction task. I offer a possible resolution to these arguments in the final suggestion for *further exploration of the dyadic advantage*.

### ***Implications of the dyadic advantage***

Human beings can form categories when working in isolation. Most previous research on “private” categorization, though, has focused on simple structures that define a small number of categories of a small number of exemplars. In the present study, I increased the complexity of the category structure and the number of categories and exemplars only slightly beyond the typical levels. Yet individual learners had a difficult time learning functions and combining them into categories, while dyadic learners accomplished this task with greater efficiency. This dyadic advantage might imply that, under time and task strictures that resemble the present study, the “public” act of referential communication plays a necessary role in the *efficient* formation of more complex “private” categories. Referential communication both distributes and consolidates the cognition of interlocutors.

In terms of psycholinguistic research, the present study offers extra-linguistic evidence that the “private” concepts interlocutors *take away* from the “public” act of referential communication differ from those formed in private. For example, dyadic learners sorted creatures into clusters that better matched the “true” category clusters and allocated attention widely enough to predict functional categories. In the absence of conversation, category learning appeared slow and laborious for individual learners. Further, the present study offers converging evidence on the benefits of collaborative problem solving and decision-making. An advantage in categorizing a novel problem might explain the oft-observed collaborative advantages in making decisions and solving problems (see, Johnson & Johnson, 2005, for review).

### ***Further exploration of the dyadic advantage***

In terms of understanding the learning advantages observed among dyadic learners, the present study was limited in three ways. Whereas theorists (*e.g.*, Chui *et al.*, 1998; Steels & Belpaeme, 2005) have claimed that communication might impose structure on an ill-defined activity and the objects, actions, and events it entails, the present study restricted category formation to one category structure with one normative classification scheme. Asking whether the dyadic advantage would persist where no normative structure exists seems an obvious follow-up question. To address this question, one might redesign the learning activity so that perceptual and latent features relate to each other and with the actions of collaborators in any of many possible ways. Then, if dyadic learners exhibit a greater tendency to form categories than individual learners one can attribute that difference to referential communication with greater confidence.

Also, by providing feedback on each of the functions entailed in a prediction rather than feedback on the category prediction as a whole, the present study does not clearly answer whether dyadic learners better learned categories or better learned each of the component functions. One can address this question with a simple modification to the present study: a dyadic condition without function-related feedback.

Finally, the present study did not control for any benefits from linguistic encoding, *per se*. Again, a simple modification to the present study—an individual condition where participants engage in private speech—could address what benefits might derive from thinking aloud. Moreover, if participants later based predictions on recordings of that private speech, one could also address whether or not performing only the beamer role requires fewer cognitive resources than performing both the spotter role and the beamer role.



## **Referential communication enhances simple-rule learning more than complex-rule learning**

### ***Summary of Results***

The results also demonstrate that the dyadic learning advantage was greater for the *simple-rule* predicted function than for the *family-resemblance* predicted function. Throughout the Function Prediction (learning) task, dyads predicted the *simple-rule* function with greater accuracy than the *family-resemblance* function. That said, the results of the posttests demonstrate that *all participants* came to exhibit this asymmetry between *simple-rule* learning and *family-resemblance* learning. They sorted more in accordance with the *simple-rule* structure. They predicted the *simple-rule* function with greater accuracy than the *family-resemblance* function, and they attended to the *simple-rule* related feature with greater likelihood than the *family-resemblance* related features when making those predictions.

### ***Explanation of the simplicity advantage***

Reference provides an explicit “rule” for an otherwise implicit judgement of how perceptual and functional features relate to one another and to the actions of interlocutors. Again as an example, the members of dyad D10.1 used “juicy-eyed” to reference the relationship between type of eyes and the nutritive “jelly” that the beamer should “juice” or extract from the creatures. Simple relationships ease the process of establishing joint reference. Dyads appear to have established reference to the simple-rules earlier in the learning process, from which point they could explicitly monitor when the referenced rule did or did not apply.

That said, the simplicity advantage observed in dyadic function learning mirrors the well-documented simplicity principle in “private” category learning—*i.e.*, human beings tend to learn simple relationships between features and their significance more easily than complex relationships (Feldman, 2003b). After all, during the posttests, individual learners (whether learning for 160 or 320 trials) also exhibited better learning of the *simple rule* than the *family-resemblance rule*. Thus, one could argue that referential communication simply increases the efficiency of learning: *i.e.*, explicit learning is simply more efficient than implicit learning.

Then again, dyadic learners learned the *family resemblance rule* better than individual learners learned the *simple rule*. *Dyadic learners predicted the family resemblance function more accurately than individual learners predicted the simple rule function during the posttest*. Dyadic-learner accuracy when predicting the *family resemblance* function ( $M = .748$ ,  $SD = .162$ ) differed significantly from the individual-learner accuracy when predicting the *simple rule* function ( $M = .646$ ,  $SD = .202$ ) —  $t_{(70)} = 2.321$ ,  $p = .023$ ). Moreover, *dyadic learners relied on family-resemblance about as much as individual learners relied on the simple-rule when sorting and when attending to features*. The jaccard similarity between the co-occurrence vectors of dyadic learners and the *family-resemblance* defined co-occurrences ( $M = .281$ ,  $SD = .217$ ) did not differ significantly from the jaccard similarity between the co-occurrence vectors of individual learners and the *simple-rule* defined co-occurrences ( $M = .241$ ,  $SD = .141$ ) —  $t_{(70)} = .914$ ,  $p = .364$ ). Also, the probability of a dyadic learner uncovering a *family-resemblance* feature ( $M = .705$ ,  $SD = .391$ ) did not differ significantly from the

probability of an individual learner uncovering a *simple rule* feature ( $M = .757$ ,  $SD = .223$ ) —  $t_{(70)} = .679$ ,  $p = .499$ .

In other words, referential communication appears to compress or simplify the *family-resemblance* structure, as well. For example, the aforementioned dyad (D10.1) referred to the prototypical destructive creature as “hungry, and carrying a knife and fork.” Such a reference abstracts away the details of surface features (*e.g.*, moving torso design, uneven tentacle length, and sharp-ended fins), rendering a simple yet concrete representation of a creature ready to attack. Variations on this economical representation likewise reduced the effort of judging creatures that differed from the prototype.

### ***Implications of the simplicity advantage***

Considered alone, the efficiency of learning simple rules might imply that structural complexity (or simplicity) constrains “public” and “private” category learning in similar ways. This finding clearly extends psycholinguistic research, where studies rarely control for structural complexity. Also, this finding reinforces existing research on category learning, offering further support for dual-system (explicit and implicit systems) theories of categorization. The apparent efficiency of establishing reference to a simple rule (*i.e.*, explicitly stating that rule) yields efficiency in learning that simple rule.

Nevertheless, the *perceived* (and, perhaps, *practical*) complexity of a concept depends in part on the shareability of the convention used in referring to the relationships between perceptual and functional features and between features and actions. For example, the use of basic-level category labels compresses the maximal number of concept-differentiating features into single word references such as *dog* or *hammer*. Certain conversations may generate equally shareable (in this case, compressed)

references to either complex or simple relationships; other conversations may generate references that complicate the simple relationships to which they refer. In either case, one cannot pre-state dyadic learning outcomes based on the category structure or its complexity alone.

The same might hold true for collaborative problem solving and decision-making, where researchers often compare outcomes from well-structured (simple) and ill-structured (complex) problems (*c.f.*, Kapur & Kinzer, 2007). Like the dyads in the present study, problem-solving groups have been found to converge on *ad hoc* conventions early in conversation (Kapur, Voiklis, & Kinzer, 2008; Voiklis, Kapur, Kinzer, & Black, 2006). The shareability of such conventions may predict the performance of problem-solving and decision-making groups.

### ***Further exploration of the simplicity advantage***

In terms of understanding the simplicity advantage, the present study was limited in two ways. First, each rule—both simple and complex—predicted different and orthogonal functions. While this design may isolate the effects of structural complexity, one might ask whether simplicity *clearly* dominates when rules of differing complexity predict the same function and the same course of action? One way to answer this question is to simply have both rules predict the same function; if most participants settle for the most frugal solution—the *simple rule*—then one might argue that simplicity constrains both “public” and “private” category learning, with greater confidence. Better yet, one might again design an activity where perceptual and latent features relate to each other and with the actions of collaborators in any of many possible ways. One could then compare what structures emerge from individual and collaborative learning contexts.

Then, if dyads extract mainly simple rules from such an ill-defined informational environment one might attribute this bias to some general simplicity advantage.

The second limitation stems from allowing dyads to invent their own referential lexicons. One might respond to the present study by asking whether and to what extent the shareability (in terms of compression) of referential conventions enhanced or hindered the learning of either the simple or complex rules. One might address this question by redesigning the present study so that a confederate of the experimenter refers to creatures using different, more or less shareable, conventions.

### **Referential communication generates conceptual homogeneity**

#### ***Summary of Results***

Finally, the results also demonstrate that dyadic learners exhibited greater conceptual homogeneity than individual learners. Dyadic partners tended to reference to the same observable features throughout the learning task. Later, each partner attended to those same features when making predictions on his or her own. Moreover, dyadic learners sorted creatures into clusters that resembled those of other dyadic learners, both within and between dyads. Post-learning sort clusters were significantly more similar between dyads than between individual learners.

#### ***Explanation of conceptual homogeneity***

Referential communication imposes shareability constraints: interlocutors try to minimize the joint cognitive effort of jointly construing an activity's various referents and their significance (Clark & Wilkes-Gibbs, 1986). Dyads appear to assure at least a

minimal level of shareability by quickly establishing *ad hoc* conventions of reference. Adherence to these conventions may reinforce attention to the features towards which conventional references point. These “public” attentional patterns may later direct “private” attention towards the same features. Moreover, similarities in referring to similar creatures may delineate a particular structure relating those creatures. Again, adherence to partner-specific conventions of reference may reinforce that structure. Conceptual homogeneity between dyads may signal a *conceptual attractor* in communication. That is, communication may push “public” conceptualizations and publicly-formed “private” conceptualizations towards a limited range of widely shareable conceptual structures.

That said, participants in the present study learned (albeit indirectly) a normative category structure. Dyadic learners learned this normative structure much better than individual learners. Between-dyad conceptual homogeneity may represent little more than normative learning of an experimenter-designed classification scheme.

Then again, dyadic learners exhibited base-level conceptual homogeneity—establishing and adhering to *ad hoc* conventions of reference—during the first block of learning trials, far earlier than they exhibited normative learning. To a certain extent, this base-level conceptual homogeneity may enhance normative learning. Specifically, *ad hoc* conventions of reference may stabilize perceptual features just enough so that dyadic learners can test hypotheses of how those perceptual features relate to functional features and how features relate to actions. This does not mean that conceptual homogeneity always leads to better learning; in fact, complete and uncritical homogeneity—groupthink—often yields cognitive myopia, leaving many hypotheses unexplored and

untested. (*c.f.*, Janis, 1982). Instead, dyadic learning may require what Voiklis *et al.* (2006) described as “tensile” intersubjectivity, where collaborators agree on the goal and parameters of the activity to which they can apply their differing skills and heuristics.

### ***Implications of conceptual homogeneity***

Group think—a dysfunctional form of conceptual homogeneity—has been widely observed in studies on collaborative work groups (*e.g.*, Janis, 1982). Psycholinguists have often observed the emergence of conceptual homogeneity—similar lexicon and similar syntax—among interlocutors and groups of interlocutors (see Branigan, Pickering, & Cleland, 2000, and Pickering & Branigan, 1999, on syntactic coordination; see Clark, 1996, for review of lexical coordination). Research on reference and category learning has provided extra-linguistic evidence (similarity in sorting and similarity in typicality judgements) of within and between dyad conceptual homogeneity (Markman & Makin, 1998; Malt & Sloman, 2004). The present study accords with these findings, but does not clearly extend them. The dyadic advantage in learning a normative structure makes it unclear whether and to what extent between-dyad conceptual homogeneity derives from some kind of conceptual attractor—a limited range of widely shareable conceptual structures—in communication.

### ***Further exploration of conceptual homogeneity***

As with the observed dyadic advantage and the observed simplicity advantage, the single normative category structure that underlay learning in the present study may limit generalizations concerning conceptual homogeneity and any in limits in the range of widely shareable conceptual structures in communication. Again, as an extension of the

present study, one would need to design an activity where perceptual and latent features relate to each other and with the actions of collaborators in any of many possible ways. Then, if pairs of interlocutors converge on similar referential conventions and/or similar ways of categorizing stimuli one can, with greater confidence, attribute that conceptual homogeneity to a conceptual attractor in communication.



## REFERENCES

- Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General*, 120(1), 3-19.
- Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, 93(3), 409-429.
- Atran, S. (2005). Adaptationism for Human Cognition: Strong, Spurious or Weak? *Mind & Language*, 20(1), 39-67.
- Atran, S., Medin, D. L., & Ross, N. (2004). Evolution and Devolution of Knowledge: A Tale of Two Biologies. *Journal of the Royal Anthropological Institute*, 10(2), 395-421.
- Austin, J. L. (1976). *How to Do Things with Words*: Oxford University Press.
- Balaban, M. T., & Waxman, S. R. (1997). Do Words Facilitate Object Categorization in 9-Month-Old Infants? *Journal of Experimental Child Psychology*, 64(1), 3-26.
- Baldwin, D. A. (1991). Infants' Contribution to the Achievement of Joint Reference. *Child Development*, 62(5), 875-890.
- Bangerter, A., & Clark, H. H. (2003). Navigating joint projects with dialogue. *Cognitive Science*, 27(2), 195-225.
- Barr, D. J., & Kronmüller, E. (2006). Conversation as a Site of Category Learning and Category Use. In A. B. Markman & B. H. Ross (Eds.), *Psychology of Learning and Motivation: Categories in Use* (Vol. 47, pp. 181-211): Elsevier.
- Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences*, 19(1), 107-154.
- Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, 11(3), 211-227.
- Barsalou, L. W., & Sewell, D. R. (1984). *Constructing representations of categories from different points of view*: Emory Cognition Project, Emory University. Document Number)
- Beller, S., & Bender, A. (2008). The Limits of Counting: Numerical Cognition Between Evolution and Culture. *Science*, 319(5860), 213-215.
- Berlin, B., Breedlove, D. E., & Raven, P. H. (1966). Folk Taxonomies and Biological Classification. *Science*, 154(3746), 273-275.
- Bloom, P. (2002). Mindreading, Communication and the Learning of Names for Things. *Mind & Language*, 17(1 & 2), 37-54.
- Blume, A., Dejong, D. V., Kim, Y. G., & Sprinkle, G. B. (1998). Experimental Evidence on the Evolution of Meaning of Messages in Sender-Receiver Games. *The American Economic Review*, 88(5), 1323-1340.
- Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, 75(2), B13-B25.

- Brennan, S. E., & Clark, H. H. (1996). Conceptual Pacts and Lexical Choice in Conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1482-1493.
- Brown, R. W. (1958). How shall a thing be called. *Psychological Review*, 65(1), 14-21.
- Bulmer, R. (1967). Why is the Cassowary Not a Bird? A Problem of Zoological Taxonomy Among the Karam of the New Guinea Highlands. *Man*, 2(1), 5-25.
- Cabrera, A., & Billman, D. (1996). Language-Driven Concept Learning: Deciphering Jabberwocky. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(2), 539-555.
- Carmichael, L., Hogan, H. P., & Walter, A. A. (1932). An experimental study of the effect of language on the reproduction of visually perceived form. *Journal of Experimental Psychology*, 15(1), 73-86.
- Cavalli-Sforza, L. L., & Feldman, M. W. (1983). Paradox of the Evolution of Communication and of Social Interactivity. *PNAS*, 80(7), 2017-2021.
- Chater, N., & Vitányi, P. M. B. (2003). Simplicity: a unifying principle in cognitive science? *Trends in Cognitive Sciences*, 7(1), 19-22.
- Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2), 121-152.
- Chiu, C. Y., Krauss, R. M., & Lau, I. Y. M. (1998). Some Cognitive Consequences of Communication. In S. R. Fussell & K. R. J. (Eds.), *Social and cognitive approaches to interpersonal communication*. Hillsdale, NJ: Erlbaum.
- Clark, A. (2006). Language, embodiment, and the cognitive niche. *Trends in Cognitive Sciences*, 10(8), 370-374.
- Clark, A., & Karmiloff-Smith, A. (1993). The Cognizer's Innards: A Psychological and Philosophical Perspective on the Development of Thought. *Mind & Language*, 8(4), 487-519.
- Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. Macwhinney (Ed.), *Mechanisms of language acquisition* (pp. 1-33). Hillsdale, NJ: Lawrence Erlbaum.
- Clark, H. H. (1972). Difficulties people have in answering the question "Where is it?". *Journal of Verbal Learning and Verbal Behavior*, 11(3), 265-277.
- Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington, DC: American Psychological Association.
- Clark, H. H., & Krych, M. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50(1), 62-81.
- Clark, H. H., & Lucy, P. (1975). Understanding what is meant from what is said: A study in conversationally conveyed requests. *Journal of Verbal Learning and Verbal Behavior*, 14(1), 56-72.
- Clark, H. H., & Marshall, C. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. Webber & I. Sag (Eds.), *Elements of discourse understanding* (pp. 10-63): Cambridge University Press.
- Clark, H. H., & Schaefer, E. (1989). Contributing to discourse. *Cognitive Science*, 13(2), 259-294.

- Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior*, 22, 245-258.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22(1), 1-39.
- Corter, J., & Gluck, M. (1992). Explaining basic categories: Feature predictability and information. *Psychological Bulletin*, 111(2), 291-303.
- Corter, J., Gluck, M., & Bower, G. (1988). *Basic levels in hierarchically structured categories*. Paper presented at the Proceedings of the 10th Annual Conference of the Cognitive Science Society.
- Crawford, V. P., & Sobel, J. (1982). Strategic Information Transmission. *Econometrica*, 50(6), 1431-1451.
- Cruse, D. (1977). The pragmatics of lexical specificity. *Journal of Linguistics*, 13(2), 153-164.
- Daniel, T. C. (1972). Nature of the effect of verbal labels on recognition memory for form. *Journal of Experimental Psychology*, 96(1), 152-157.
- Davidson, D. (1975). Thought and Talk. In H. Geirsson & M. Losonsky (Eds.), *Readings in Language and Mind: An Interdisciplinary Reader* (pp. 155-170). Cambridge, MA: Blackwell Publishing.
- Deacon, T. W. (2000). Evolutionary perspectives on language and brain plasticity. *Journal of Communication Disorders*, 33(4), 273-291.
- Dehaene, S. (2001). Precis of The Number Sense. *Mind & Language*, 16(1), 16-36.
- Deutsch, M. (1949). A Theory of Co-operation and Competition. *Human Relations*, 2(2), 129-152.
- Deutsch, W., & Pechmann, T. (1982). Social interaction and the development of definite descriptions. *Cognition*, 11(2), 159-184.
- Diamond, J. (1966). Zoological Classification System of a Primitive People. *Science*, 151(3714), 1102-1104.
- Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature*, 407(6804), 630-633.
- Feldman, J. (2003a). A catalog of Boolean concepts. *Journal of Mathematical Psychology*, 47(1), 75-89.
- Feldman, J. (2003b). The simplicity principle in human concept learning. *Current Directions in Psychological Science*, 12(6), 227-232.
- Feldman, J. (2006). An algebra of human concept learning. *Journal of Mathematical Psychology*, 50(4), 339-368.
- Francik, E. P., & Clark, H. H. (1985). How to make requests that overcome obstacles to compliance. *Journal of Memory and Language*, 24(5), 560-568.
- Freyd, J. (1983). Shareability: The social psychology of epistemology. *Cognitive Science*, 7(3), 191-210.
- Furnas, G. W., Landauer, T. K., Gomez, L. M., & Dumais, S. T. (1987). The vocabulary problem in human-system communication. *Communications of the ACM*, 30(11), 964-971.
- Fussell, S. R., & Krauss, R. M. (1989). The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of Experimental Social Psychology*, 25(3), 203-219.

- Gallese, V., & Lakoff, G. (2005). The Brain's concepts: the role of the Sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22(3), 455-479.
- Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27(2), 181-218.
- Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistic conventions. *Cognition*, 53(3), 181-215.
- Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. W. (1998). Training 'greeble' experts: a framework for studying expert object recognition processes. *Vision Research*, 38(15-16), 2401-2428.
- Gelman, R., & Gallistel, C. R. (2004). Language and the origin of numerical concepts. *Science*, 306(5695), 441-443.
- Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, 23(3), 183-209.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103(4), 650-669.
- Gluck, M. A., & Corter, J. E. (1985). *Information, uncertainty, and the utility of categories*. Paper presented at the Proceedings of the Seventh Annual Conference of the Cognitive Science Society, Hillsdale, NJ.
- Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do People Solve the ``Weather Prediction" Task?: Individual Variability in Strategies for Probabilistic Category Learning. *Learning & Memory*, 9(6), 408-418.
- Goodenough, W. H. (1965). Yankee Kinship Terminology: A Problem in Componential Analysis. *American Anthropologist*, 67(5, Part 2: Formal Semantic Analysis), 259-287.
- Gordon, P. (2004). Numerical Cognition Without Words: Evidence from Amazonia. *Science*, 306(5695), 496-499.
- Gosselin, F., & Schyns, P. G. (2001). Why do we SLIP to the basic level? Computational constraints and their implementation. *Psychological Review*, 108(4), 735-758.
- Gould, S. J., & Lewontin, R. C. (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 205(1161), 581-598.
- Grice, H. P. (1957). Meaning. *The Philosophical Review*, 66(3), 377-388.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics 3: Speech acts* (pp. 41-58). New York: Academic Press.
- Hamza, W., Bakis, R., Eide, E. M., Picheny, M. A., & Pitrelli, J. F. (2004). *The IBM Expressive Speech Synthesis System*. Paper presented at the Proceedings of the International Conference on Spoken Language Processing.
- Hauser, M. D. (1997). Artifactual kinds and functional design features: what a primate understands without language. *Cognition*, 64(3), 285-308.
- Hauser, M. D. (2000). What do animals think about numbers. *American Scientist*, 88(2).
- Hauser, M. D. (2005). Our chimpanzee mind. *Nature*, 437(7055), 60-63.
- Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 267(1445), 829-833.

- Hauser, M. D., Macneilage, P., & Ware, M. (1996). Numerical representations in primates. *PNAS*, 93(4), 1514-1517.
- Higgins, E. T. (1981). The "communication game": Implications for social cognition and persuasion. In E. T. Higgins, C. P. Herman & M. P. Zanna (Eds.), *Social Cognition: The Ontario Symposium, Vol. 1* (pp. 343-392). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hoffmann, J., & Ziessler, C. (1983). Objektidentifikation in kunstlichen Begriffshierarchien [Object identification in artificial concept hierarchies]. *Zeitschrift fur Psychologie*, 191(2), 135-167.
- Hood, B. M., Hauser, M. D., Anderson, L., & Santos, L. (1999). Gravity biases in a non-human primate? *Developmental Science*, 2(1), 35-41.
- Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59(1), 91-117.
- Hulstijn, J., & Maudet, N. (2006). Uptake and joint action. *Cognitive Systems Research*, 7(2-3), 175-191.
- Hunn, E. S. (1977). *Tzeltal folk zoology: the classification of discontinuities in nature*: Academic Press.
- Hutchins, E., & Hazlehurst, B. (1992). Learning in the Cultural Process. In C. G. Langton, C. Taylor, D. J. Farmer & S. Rasmussen (Eds.), *Artificial Life II* (pp. 689-706). Redwood City, CA: Addison-Wesley.
- Isaacs, E. A., & Clark, H. H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology: General*, 116(1), 26-37.
- Janis, I. L. (1982). *Groupthink: psychological studies of policy decisions and fiascoes*: Houghton Mifflin.
- Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and Competition: Theory and Research*. Edina, MN: Interaction Book Company.
- Johnson, D. W. and Johnson, R. T. (2005). New developments in social interdependence theory. *Genetic, Social, and General Psychology Monographs*, 131(4):285-358.
- Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology*, 16(2), 243-275.
- Jones, G. (1983). Identifying basic categories. *Psychological Bulletin*, 94(3), 423-428.
- Kaplan, F., & Hafner, V. V. (2006). The challenges of joint attention. *Interaction Studies*, 7(2), 135-169.
- Kapur, M., & Kinzer, C. (2007). Examining the effect of problem type in a synchronous computer-supported collaborative learning (CSCL) environment. *Educational Technology, Research and Development*, 55(5), 439-459.
- Kapur, M., Voiklis, J., & Kinzer, C. (2008). Sensitivities to early exchange in synchronous computer-supported collaborative learning (CSCL) groups. *Computers & Education*, 51(1), 54-66.
- Katz, P. A. (1963). Effects of labels on children's perception and discrimination learning. *Journal of experimental psychology*, 66, 423-428.
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking Perspective in Conversation: The Role of Mutual Knowledge in Comprehension. *Psychological Science*, 11(1), 32-38.

- Keysar, B., Barr, D. J., Balin, J. A., & Paek, T. S. (1998). Definite Reference and Mutual Knowledge: Process Models of Common Ground in Comprehension. *Journal of Memory and Language*, 39(1), 1-20.
- Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, 89(1), 25-41.
- Kotovsky, L., & Gentner, D. (1996). Comparison and Categorization in the Development of Relational Similarity. *Child Development*, 67(6), 2797-2822.
- Krauss, R. M., & Fussell, S. R. (1991). Perspective-taking in communication: Representations of others' knowledge in reference. *Social Cognition*, 9(1), 2-24.
- Krauss, R. M., & Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction: A preliminary study. *Psychonomic Science*, 1, 113-114.
- Kronmüller, E., & Barr, D. J. (2007). Perspective-free pragmatics: Broken precedents and the recovery-from-preemption hypothesis. *Journal of Memory and Language*, 56(3), 436-455.
- Landau, B., & Shipley, E. (2001). Labelling patterns and object naming. *Developmental science*, 4(1), 109-118.
- Lewis, D. (1969). *Convention: a philosophical study*: Harvard University Press.
- Lewis, D. K. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8(1), 339-359.
- Lieberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, 4(5), 187-196.
- Lin, E. L., & Murphy, G. L. (1997). Effects of background knowledge on object categorization and part detection. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1153-1169.
- Livingston, K. R., & Andrews, J. K. (1995). On the Interaction of Prior Knowledge and Stimulus Structure in Category Learning. *The Quarterly Journal of Experimental Psychology Section A*, 48(1), 208-236.
- Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, 50(4), 315-353.
- Love, B. C. (2002). Comparing supervised and unsupervised category learning. *Psychonomic Bulletin & Review*, 9, 829-835.
- Love, B. C. (2003). The multifaceted nature of unsupervised category learning. *Psychonomic Bulletin & Review*, 10, 190-197.
- Lupyan, G. (2008). From Chair To "Chair:" A Representational Shift Account Of Object Labeling Effects On Memory. *Journal of Experimental Psychology: General*, 137(2), 348-369.
- Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language Is Not Just for Talking: Redundant Labels Facilitate Learning of Novel Categories. *Psychological Science*, 18(12), 1077-1083.
- Maclin, K. (2002). The effects of exemplar and prototype descriptors on verbal overshadowing. *Applied Cognitive Psychology*, 16(8), 929-936.
- Malt, B. C., & Sloman, S. A. (2004). Conversation and convention: Enduring influences on name choice for common objects. *Memory and Cognition*, 32(8), 1346-1354.

- Malt, B. C., Sloman, S. A., & Gennari, S. (2003a). Speaking vs. thinking about objects and actions. In D. Gentner & G. S. Meadow (Eds.), *Language in Mind: Advances in the Study of Language and Thought* (pp. 81-111). Cambridge, MA: MIT Press.
- Malt, B. C., Sloman, S. A., & Gennari, S. (2003b). Universality and language specificity in object naming. *Journal of Memory and Language*, 49(1), 20-42.
- Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus Naming: Similarity and the Linguistic Categorization of Artifacts. *Journal of Memory and Language*, 40(2), 230-262.
- Markman, A. B., & Makin, V. S. (1998). Referential communication and category acquisition. *Journal of Experimental Psychology: General*, 127(4), 331-354.
- Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin*, 129(4), 592-613.
- Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20(2), 121-157.
- Matsuka, T., & Corter, J. E. (in press). Observed attention allocation processes in category learning. *The Quarterly Journal of Experimental Psychology*.
- Medin, D. L., & Atran, S. (2004). The native mind: Biological categorization and reasoning in development and across cultures. *Psychological Review*, 111(4), 960-983.
- Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology*, 19(2), 242-279.
- Melcher, J., & Schooler, J. (1996). The Misremembrance of Wines Past: Verbal and Perceptual Expertise Differentially Mediate Verbal Overshadowing of Taste Memory. *Journal of Memory and Language*, 35(2), 231-245.
- Meltzoff, & Andrew, N. (2007). Like me: a foundation for social cognition. *Developmental Science*, 10(1), 126-134.
- Mervis, C., & Crisafi, M. (1982). Order of Acquisition of Subordinate-, Basic-, and Superordinate-Level Categories. *Child Development*, 53(1), 258-266.
- Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, 49(2), 201-213.
- Minda, J. P., & Ross, B. H. (2004). Learning categories by making predictions: An investigation of indirect category learning. *Memory & Cognition*, 32.
- Monos, C. M. (1997). *Learning and use of hierarchically structured categories in a social context*.
- Morris, M. W., & Murphy, G. L. (1990). Converging operations on a basic level in event taxonomies. *Memory & Cognition*, 18(4), 407-418.
- Moyer, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception and Psychophysics*, 13(2), 180-184.
- Murphy, G. L., & Brownell, H. H. (1985). Language and thought: Aspects of a cognitive theory of semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(1), 70-84.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92(3), 289-316.

- Murphy, G. L., & Smith, E. (1982). Basic-level superiority in picture categorization. *Journal of Verbal Learning and Verbal Behavior*, 21(1), 1-20.
- Murphy, G. L., & Wright, J. C. (1984). Changes in conceptual structure with expertise: Differences between real-world experts and novices. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(1), 144-155.
- Nosofsky, R. M., Gluck, M. A., Palmieri, T. J., Mckinley, S. C., & Glauthier, P. (1994). Comparing models of rule-based classificatoin learning: a replication and extension of Shepard, Hovland, and Jenkins(1961). *Memory & cognition*, 22(3), 352-369.
- Paivio, A. (1978). Mental comparisons involving abstract attributes. *Memory & Cognition*, 6, 199-208.
- Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics*, 27, 89-110.
- Pfeifer, R., & Bongard, J. C. (2006). *How the Body Shapes the Way We Think: A New View of Intelligence*. Cambridge, MA: The MIT Press.
- Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and Approximate Arithmetic in an Amazonian Indigene Group. *Science*, 306(5695), 499-503.
- Pickering, M. J., & Branigan, H. P. (1999). Syntactic priming in language production. *Trends in Cognitive Sciences*, 3(4), 136-141.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(02), 169-190.
- Pickering, M. J., & Garrod, S. (2006). Alignment as the Basis for Successful Communication. *Research on Language & Computation*, 4(2), 203-228.
- Pietarinen, A. V. (2006). The evolution of semantics and language-games for meaning. *Interaction Studies*, 7(1), 79-104.
- Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships Between Perception and Action* (pp. 167-201). Berlin: Springer-Verlag.
- Ratneshwar, S., Barsalou, L. W., Pechmann, C., & Moore, M. (2001). Goal-Derived Categories: The Role of Personal and Situational Goals in Category Representations. *Journal of Consumer Psychology*, 10(3), 147-157.
- Rifkin, A. (1985). Evidence for a basic level in event taxonomies. *Memory & Cognition*, 13(6), 538-556.
- Romney, K., & D'Andrade, R. (1964). Cognitive Aspects of English Kin Terms. *American Anthropologist*, 66(3, Part 2: Transcultural Studies in Cognition), 146-170.
- Romney, K. A., Boyd, J. P., Moore, C. C., Batchelder, W. H., & Brazill, T. J. (1996). Culture as shared cognitive representations. *PNAS*, 93(10), 4699-4705.
- Rosch, E., Mervis, C., Gray, W., Johnson, D., & Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8(3), 382-439.
- Ross, B. H. (1997). The Use of Categories Affects Classification. *Journal of Memory and Language*, 37(2), 240-267.
- Ross, B. H. (1999). Postclassification category use: The effects of learning to use categories after learning to classify. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(3), 743-757.
- Russell, B. (1905). On Denoting. *Mind*, 14(56), 479-493.



- Sacks, H., Schegloff, E., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50(4), 696-735.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, 16(2), 235-239.
- Schelling, T. C. (1960). *The Strategy of Conflict*. Harvard University Press.
- Schooler, J., & Engstler-Schooler, T. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, 22(1), 36-71.
- Schvaneveldt, R. W., Durso, F. T., Goldsmith, T. E., Breen, T. J., Cooke, N. M., Tucker, R. G., et al. (1985). Measuring the structure of expertise. *International journal of man-machine studies*, 23(6), 699-728.
- Shepard, R. (1992). The perceptual organization of colors: An adaptation to regularities of the terrestrial world. In J. Barkow, L. Cosmides & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 495-532).
- Shepard, R., Hovland, C., & Jenkins, H. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, 75(13), 1-42.
- Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, 69(1), 99-118.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63(2), 129-138.
- Smith, L., Jones, S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object Name Learning Provides On-the-Job Training for Attention. *Psychological Science*, 13(1), 13-19.
- Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: a dumb attentional mechanism? *Cognition*, 60(2), 143-171.
- Stalnaker, R. C. (1978). Assertion. In P. Cole (Ed.), *Syntax and Semantics 9: Pragmatics* (pp. 315-332). New York: Academic Press.
- Stalnaker, R. C. (2002). Common Ground. *Linguistics and Philosophy*, 25(5-6), 701-721.
- Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, 23(3), 457-482.
- Thompson, R. K. R., Oden, D. L., & Boysen, S. T. (1997). Language-naïve chimpanzees (Pan troglodytes) judge relations between relations in a conceptual matching-to-sample task. *Journal of Experimental Psychology: Animal Behavior Processes*, 23(1), 31-43.
- Todd, P., & Gigerenzer, G. (2007). Environments That Make Us Smart: Ecological Rationality. *Current Directions in Psychological Science*, 16(3), 167-171.
- Todd, P. M., & Gigerenzer, G. (2001). Précis of Simple heuristics that make us smart. *Behavioral and Brain Sciences*, 23(05), 727-741.
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5), 675-691.
- Tomasello, M., & Merriman, W. E. (1995). *Beyond Names for Things: Young Children's Acquisition of Verbs*. Lawrence Erlbaum.

- Tooby, J., & Cosmides, L. (1989). Evolutionary psychology and the generation of culture, part I : Theoretical considerations. *Ethology and Sociobiology*, 10(1-3), 29-49.
- Tversky, B. (2005). Visuospatial Reasoning. In K. J. Holyoak & R. G. Morrison Jr (Eds.), *The Cambridge Handbook of Thinking and Reasoning*. New York: Cambridge University Press.
- Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of experimental psychology: General*, 113(2), 169-197.
- Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, 94(1), 3-15.
- Voiklis, J., Kapur, M., Kinzer, C., & Black, J. (2006). *An Emergentist Account of Collective Cognition in Collaborative Problem Solving*. Paper presented at the Proceedings of the 28th Annual Conference of the Cognitive Science Society.
- Vygotsky, L. S. (1986 [1962]). *Thought and Language* (A. Kozulin, Trans.). Cambridge, MA: MIT Press.
- Wagner, W., & Wagner, G. P. (2004). Examining the Modularity Concept in Evolutionary Psychology: The Level of Genes, Mind, and Culture. *Journal of Cultural and Evolutionary Psychology*, 1(3), 135-165.
- Waxman, S. R., & Booth, A. E. (2001). Seeing Pink Elephants: Fourteen-Month-Olds Interpretations of Novel Nouns and Adjectives. *Cognitive Psychology*, 43(3), 217-242.
- Waxman, S. R., & Markow, D. B. (1995). Words as Invitations to Form Categories: Evidence from 12- to 13-Month-Old Infants. *Cognitive Psychology*, 29(3), 257-302.
- Whorf, B. L. (1956). *Language, thought and reality: Selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.
- Wiley, J., & Jolly, C. (2003). *When two heads are better than one expert*. Paper presented at the Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society.
- Wilkes-Gibbs, D., & Clark, H. (1992). Coordinating beliefs in conversation. *Journal of Memory and Language*, 31(2), 183-194.
- Wittgenstein, L. (2001 [1958]). *Philosophical Investigations: 50th Anniversary Commemorative Edition* (G. E. M. Anscombe & E. Anscombe, Trans.): Blackwell Publishing, Incorporated.
- Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, 85(3), 223-250.
- Yamauchi, T., & Markman, A. B. (1998). Category Learning by Inference and Classification. *Journal of Memory and Language*, 39(1), 124-148.
- Yamauchi, T., & Markman, A. B. (2000a). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 776-795.
- Yamauchi, T., & Markman, A. B. (2000b). Learning categories composed of varying instances: the effect of classification, inference, and structural alignment. *Memory & Cognition*, 28(1), 64-78.
- Zajonc, R. B. (1965). Social Facilitation. *Science*, 149(3681), 269-274.

## APPENDIX A

### Descriptive and Point-Based Feedback

Table A1. Descriptive Feedback for the Function Prediction and Attention Allocation Tasks.

Participant Response	Functional Significance of Stimulus			
	No Functional Significance	Nutritive Only	Destructive Only	Nutritive & Destructive
No Functional Significance	<i>Time and energy conserved</i>	<i>No jelly extracted</i>	<i>Life support damaged</i>	<i>Life support damaged, and no jelly extracted</i>
Nutritive Only	<i>Tractor beam wasted</i>	<i>Jelly extracted</i>	<i>Life support damaged, and tractor beam wasted</i>	<i>Life support damaged, some jelly extracted</i>
Destructive Only	<i>Stun beam wasted</i>	<i>No jelly extracted, stun beam wasted</i>	<i>Life support protected</i>	<i>Life support protected, but no jelly extracted</i>
Nutritive & Destructive	<i>Tractor beam and stun beam both wasted</i>	<i>Stun beam wasted, some jelly extracted</i>	<i>Life support protected, but tractor beam wasted</i>	<i>Life support protected, and jelly extracted</i>
No Response	<i>Too slow, are you alert?</i>	<i>Too slow, no jelly extracted</i>	<i>Too slow, life support damaged</i>	<i>Too slow, life support damaged and no jelly extracted</i>

Table A2. *Energy Meter*: Rewards and Penalties for the Function Prediction and Attention Allocation Tasks.

	<b>Functional Significance of Stimulus</b>			
<b>Participant Response</b>	No Functional Significance	Nutritive Only	Destructive Only	Nutritive & Destructive
No Functional Significance	10	-5	-10	-15
Nutritive Only	-5	10	-15	-10
Destructive Only	-5	-10	10	-5
Nutritive & Destructive	-10	5	-5	10
No Response	-10	-10	-10	-10

## APPENDIX B

### Examples of Stimuli

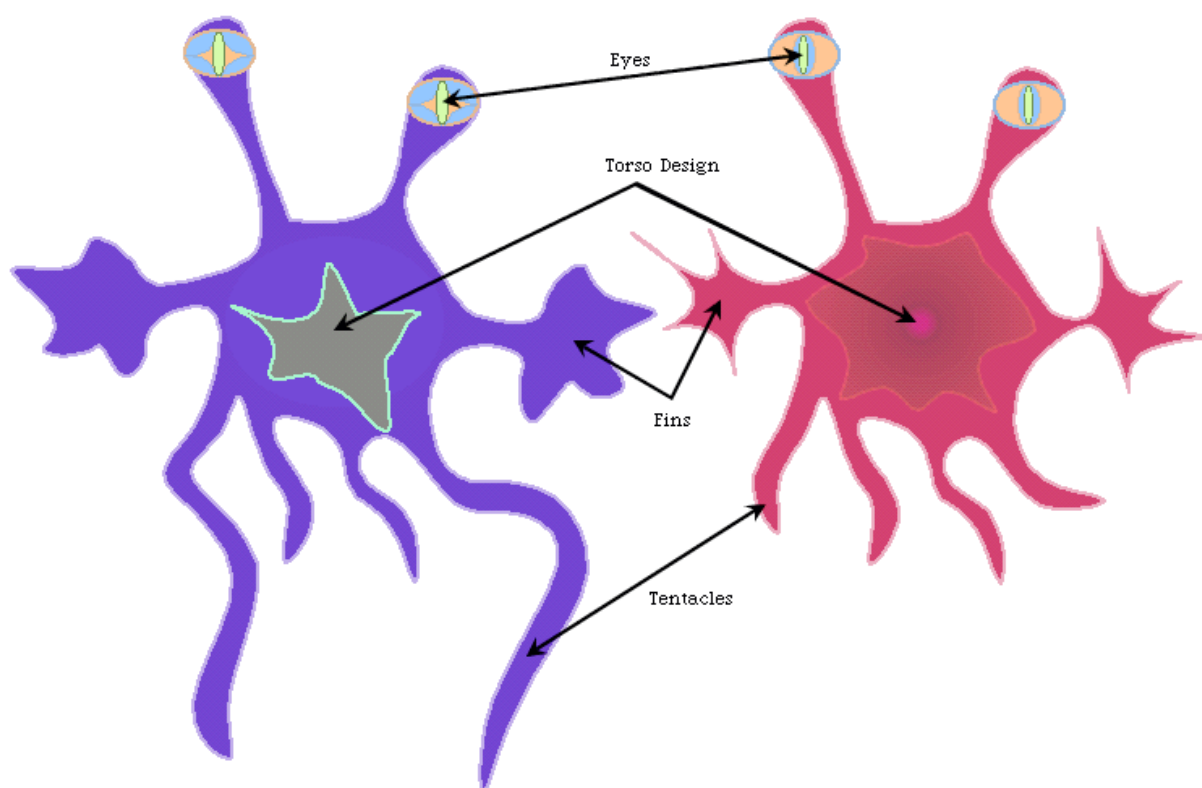


Figure B1. Examples of the binary feature values that instantiate the stimuli and the perceptual structure of the functional categories.

## APPENDIX C

### Instructions for Tasks and Task Interfaces

**Your Mission...**

You have volunteered for a three-year mission on planet Zaff. The journey from Earth to Zaff will deplete most of your food & fuel supplies. Luckily, certain creatures on planet Zaff produce a highly nutritious jelly that can also power your spaceship's bio-fuel reactor. Use a Tractor Beam <Ctrl-space> to capture these creatures. Other creatures do not produce jelly, but will sabotage your life support systems if not frozen with a Stun Beam <Option-space>. Be aware that some jelly-producing creatures will also sabotage your life support if not stunned before capturing them <Ctrl-Option-space>. Still other creatures do nothing. Shoo these creatures away with a harmless Flasher Beam <space>.

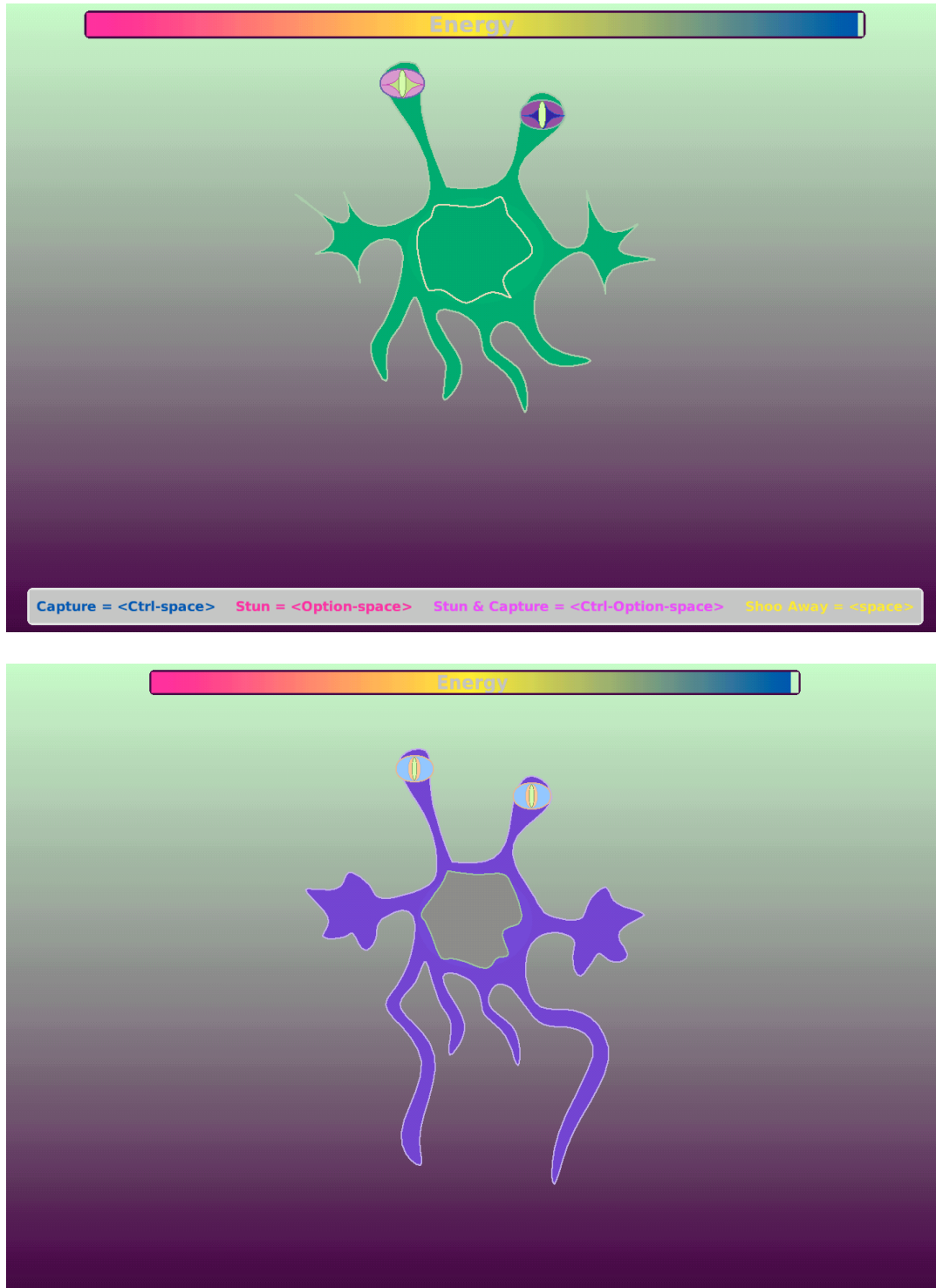
Remember: using the wrong Beam wastes both food and fuel and will leave you less able to survive your mission.

Use these training materials to learn both how to tell apart the different kinds of creatures on planet Zaff and which Beam to use on each kind of creature.

Good luck!

**Click anywhere to proceed to the next task**

Figure C1. Describes the scenario for the game tasks and provides instruction on how to execute function-prediction related actions.



For this training exercise, you will try to figure out what each different kind of Zaffian creature does -- produce jelly, sabotage your life support, both, or neither -- and what action to take -- stun, capture, stun then capture, or shoo away that kind of creature.

During the training, creatures will appear in your training console one at a time. You have fifteen seconds to decide what beam(s) to use: Capture = <Ctrl-space>, Stun = <Option-space>, Stun & Capture = <Ctrl-Option-space>, and Shoo = <space>.

Keep an eye on the energy gauge; it let's you know your total remaining food & fuel.

**Click anywhere to begin training task**

For this training exercise, you and a partner will try to figure out what each different kind of Zaffian creature does -- produce jelly, sabotage your life support, both, or neither -- and what action to take -- stun, capture, stun then capture, or shoo away that kind of creature.

When the training starts, a creature will appear either in your training console or that of your partner. The trainee (the spotter) who can see the creature has ten seconds to describe it to the other trainee (the beamer), who then has five seconds to decide what beam(s) to use: Capture = <Ctrl-space>, Stun = <Option-space>, Stun & Capture = <Ctrl-Option-space>, and Shoo = <space>. (Note: the spotter should NOT tell the beamer what the creature does nor what beam(s) to use). A new creature will then appear in the beamer's training console. In this way, you and your partner will take turns working as spotter and beamer.

Remember to keep an eye on the energy gauge; it let's you know your total remaining food & fuel.

# Partner: Not Ready

**Click anywhere to alert your partner when YOU are ready**

Figure C3 A & B. Function Prediction Task Instructions. A (top) = Single Player, and B (bottom) = Two Player.



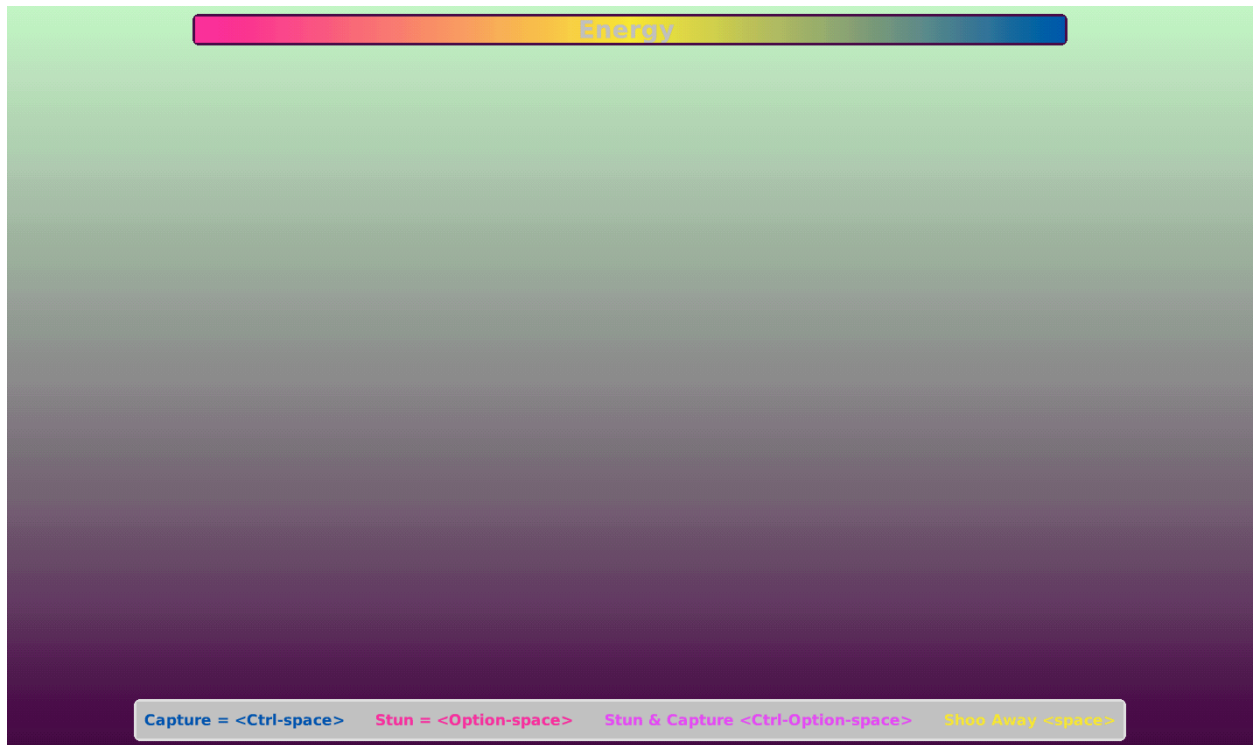


Figure C4. Two-Player Beamer Function Prediction (FP) Task Interface.

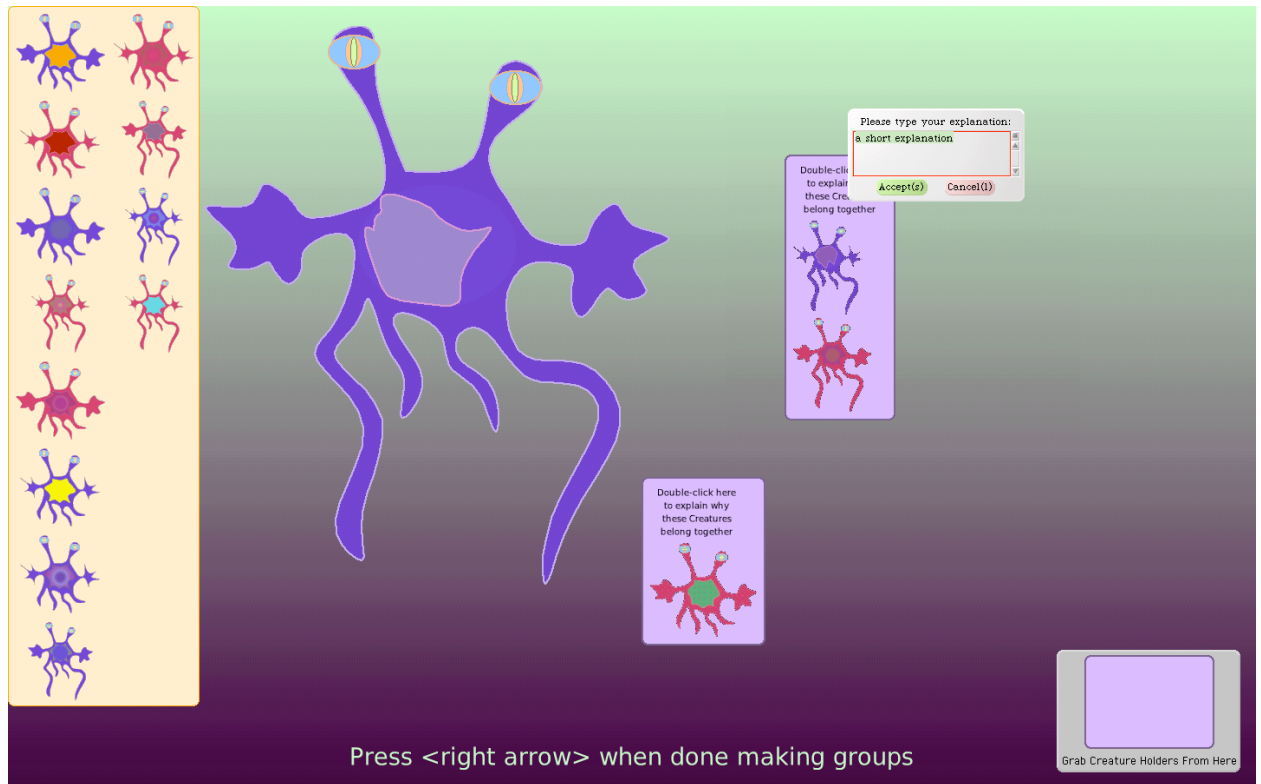


Figure C5. Interface for the PRE-sort and POST-sort tasks.

**Welcome Volunteers...**

Before we reveal your secret mission, we want you to look at different "alien" creatures and place the creatures that you think belong together into the same Creature Holder (violet-colored rectangle).

To make a group of creatures, grab a Creature Holder from the dispenser located in the lower right corner and drag it onto the Creature Screen. Then, grab a creature from the Creature Source (peach-colored rectangle) located in the upper left corner and drag it onto the screen. Once you've decided to which group the creature belongs, drag the creature into that group's Creature Holder. Make as many or as few groups as you need to use up all the creatures in the Creature Source. Use one Creature Holder per group.

For each group, double-click the explanation button and use the dialog box to type why you think the creatures in the group belong together.

**Click anywhere to begin training task**

For this training exercise, you will again make groups of Zaffian creatures by placing the creatures that you think belong together into the same Creature Holder (violet-colored rectangle).

To make a group of creatures, grab a Creature Holder from the dispenser located in the lower right corner and drag it onto the Creature Screen. Then, grab a creature from the Creature Source (peach-colored rectangle) located in the upper left corner and drag it onto the screen. Once you've decided to which group the creature belongs, drag the creature into that group's Creature Holder. Make as many or as few groups as you need to use up all the creatures in the Creature Source. Use one Creature Holder per group.

For each group, double-click the explanation button and use the dialog box to type why you think the creatures in the group belong together.

**Click anywhere to begin training task**

Figure C6 A & B. Instructions for the PRE-sort (top) and POST-sort tasks (bottom).

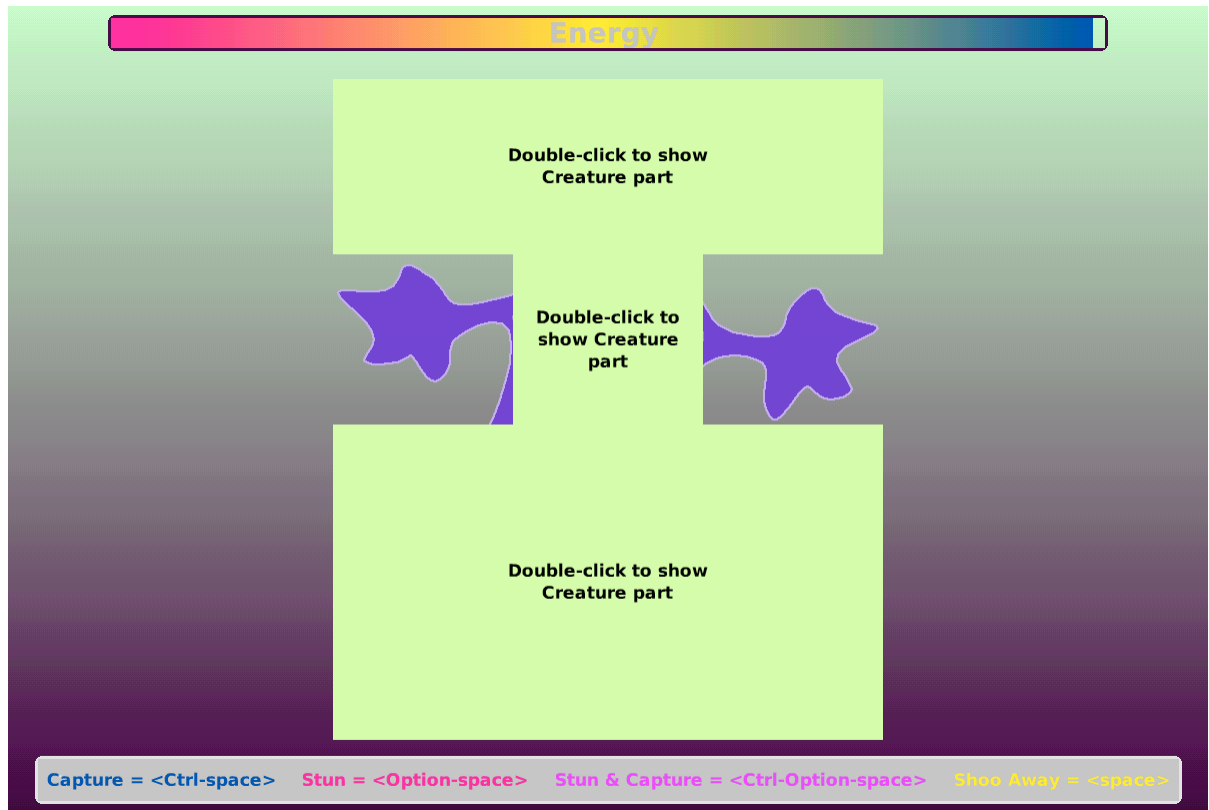


Figure C7. Attention Allocation Task Interface.

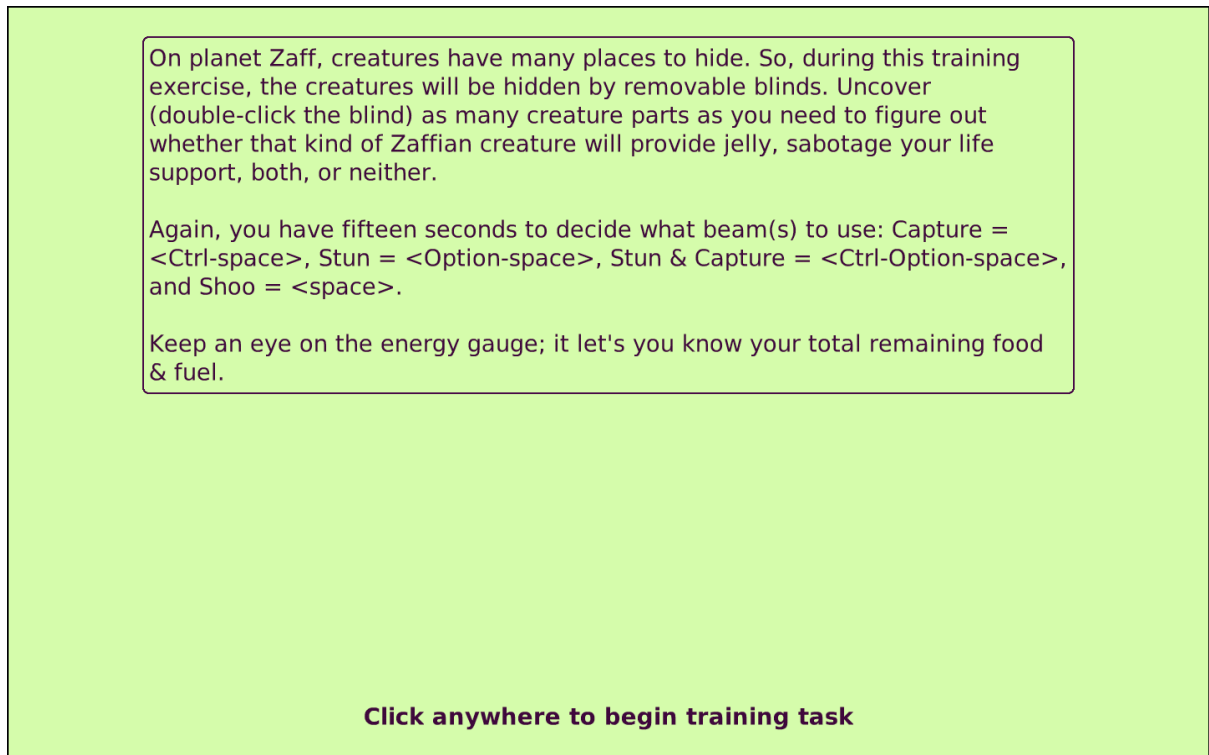


Figure C8. Instructions for Attention Allocation Task, including how to execute task-related action.

