When we hear, “these are kiwanos”, while being shown a set of objects each displaying a novel combination of yellowness, spikiness, and ovalness, how do these regularities factor into our hypotheses of what the word ‘kiwano’ means? One possibility is that we acquire a new mental term, KIWANO, by conjoining the representations YELLOW, SPIKEY, and OVAL such that the observed regularities are directly encoded as part of the meaning of the new mental term. Perhaps over time and multiple learning instances, critical information is differentiated from noise, and we come to represent KIWANO as a set of probability distributions over the most typical features. Variants of this descriptivist approach include classical definition theories (Locke, 1690; Hume, 1748), prototype theories (Rosch, 1978; Barsalou, 1987, 1999; Prinz 2002), and exemplar theories (Smith & Medin 1981). Another possibility, however, is that observed regularities are not themselves constitutive of meanings, but instead serve only as the basis for positing hidden/latent properties as causal explanations for the observed regularities. On such essentialist accounts (Gelman, 2003; Keil, 1989; Carey, 1985), the concept KIWANO may be simple in representational structure even if it was acquired from complex sets of observed similarities.

This paper fleshes out our favored essentialist account (Oved, 2009; 2014), which is made precise with a computational implementation in the spirit of Bayesian Cognitive Science (following Griffiths, Kemp, & Tenenbaum, 2008; Xu & Tenenbaum, 2007; Gopnik & Tenenbaum, 2007; Kemp, Perfors, & Tenenbaum, 2007). On this approach, many lexical concepts for natural kinds, like KIWANO, APPLE, ELEPHANT, and WATER, are mental names for properties that are posited as explanations for observed similarities and differences in perceptible properties, and are stored as values of latent variables in a Bayesian network (Oved & Fasel, 2011). After describing this essentialist account we present two experiments measuring the use of generic utterances (typified by the form ‘Fs are G’ in contrast with ‘All/most (of the) Fs are G’) to support this account over descriptivism. We conclude with a third, pilot experiment and a discussion of lexical concepts for artifact and social kinds.

By many accounts, generics are a linguistic tool for expressing lawlike relationships between kinds and their properties (Carlson, 1977; Prasada, 2000; Prasada & Dillingham, 2006, 2009; Prasada, Khemlani, Leslie & Glucksberg, 2013; Gelman, 2003; Gelman & Bloom, 2007). In two experiments we compared descriptivist approaches with our essentialist approach by exploring the conditions under which people use generics to describe an observed regularity. We introduced our participants (adult native-English speakers on Mechanical Turk) to three novel creatures from a distant galaxy by presenting shape outlines and labels, ‘toma’, ‘pimwit’, and ‘kirbo’. We then showed an image of a planet with strong correlations between creature-kind and color –all tomas were black, all pimwits were blue, and all kirbos were red– and asked the participants to describe the colors of the creatures. Participants who received information clearly indicating that the regularity was lawlike (e.g., fur colors are determined by blood type) used significantly more generics than participants who received information indicating that the regularity was accidental (e.g., fur colors are determined by the colored pool the creature bathed in that day). When simply shown the regularity between creature-kind and fur color, without any information about what determines the fur color, participants used as many generic utterances as when they were explicitly told that the connection was lawlike. This first differential use of generics suggests that generics are indeed a tool for expressing lawlike connections, and the second differential use suggests that by default, at least when it comes to creature-kinds and colors, humans interpret regularities as lawlike. While this finding is predicted by our essentialist account, descriptivist theories of concepts have no framework for accounting for this cognitive difference between lawlike and accidental correlations. We are currently piloting a version of these experiments to carry out on preschool children and future versions will use the paradigm to explore artifact and social kinds.