Evolutionary Psychology

www.epjournal.net – 2013. 11(3): 550-572

Original Article

The Origins of Pedagogy: Developmental and Evolutionary Perspectives

Amy E. Skerry, Department of Psychology, Harvard University, Cambridge, USA. Email: amy.skerry@gmail.com (Corresponding author).

Enoch Lambert, Department of Philosophy, Harvard University, Cambridge, USA.

Lindsey J. Powell, Department of Psychology, Harvard University, Cambridge, USA.

Katherine McAuliffe, Department of Human Evolutionary Biology, Harvard University, Cambridge, USA.

Abstract: The question of whether and how information is actively transferred from knowledgeable to ignorant individuals has received much attention in psychology and evolutionary biology. Research in these fields has proceeded largely independently, with studies of nonhuman animals focusing on knowledgeable individuals and whether or not they meet a functional definition of teaching, while studies of children focus on the learner's assumptions and inferences. We argue that a comprehensive theory of teaching will benefit from integrating perspectives and empirical phenomena from evolutionary and developmental disciplines. In this review, we identify cases of seemingly purposeful information transfer (i.e. teaching) in human and nonhuman animals, discuss what is known about the cognitive processes that support teaching in different species, and highlight ways in which each discipline might be informed by extant theories and empirical tools from the other.

Keywords: pedagogy, teaching, social learning, comparative cognition

Introduction

One of the primary challenges for a young organism is to learn about the world around it. Often, knowledge is available in the minds of conspecifics that would be risky, inefficient, or even impossible to derive through individual exploration of the environment. Thus, conspecifics can be rich sources of relevant information for a learner. Indeed, decades of research on humans and other species attest to the centrality of social learning for the acquisition of adaptive information and skills (see Galef, 1976; Hoppitt and Laland, 2008; Rendell et al., 2011). Observing another's actions can lead a naïve organism to focus attention on relevant environmental features (stimulus enhancement), to identify causal contingencies (affordance learning), or to model the actions or goals of another (imitation/emulation) (see Galef and Giraldeau, 2001; Hoppitt and Laland, 2008). In these cases, the learner may exploit information made available by actions the knowledgeable actor performs for the actor's own immediate benefit, without regard to the audience (Danchin, Giraldeau, Valone, and Wagner, 2004). Such passive forms of social learning are observed across diverse taxa (e.g., in fish: Laland, Atton, and Webster, 2011; insects: Leadbeater and Chittka, 2007; birds: Slagsvold and Wiebe, 2011; primates: Whiten, 2011; and other mammals: Thornton and Clutton-Brock, 2011).

However, some information may be more effectively transmitted with active involvement of the knowledgeable individual (Boyd and Richardson, 1985). In teaching, also referred to as pedagogy, a knowledgeable individual facilitates learning by altering its behavior in the presence of a naïve observer. Teaching is widespread amongst humans (Strauss, Ziv, and Stein, 2002), and according to one recent proposal (hereafter, the "Natural Pedagogy" or "NP" hypothesis), human teaching relies on dedicated mechanisms that evolved for the purpose of social transmission (Csibra and Gergely, 2009, 2011), and are unique to our species (Csibra, 2007)[1]. Alternatively, the ability to learn effectively from teachers could be acquired over the course of ontogeny via domain-general learning mechanisms, without requiring genetic adaptations for social transmission[2]. To begin to distinguish between these possibilities, recent experiments have explored how social-communicative contexts modulate learning early in human development, generating an intriguing set of empirical phenomena. However, the representations that support differential learning in teaching contexts remain poorly understood, and the extent to which there is continuity in these processes across species is unknown.

In parallel to this developmental literature, there has emerged a growing body of research on teaching in nonhuman animals (for reviews, see Caro and Hauser, 1992; Thornton and Raihani, 2008). This work is not only relevant for evaluating the claim that the form of pedagogy observed in humans is species-specific (Csibra, 2007), but also for understanding the selection pressures that might have given rise to teaching, and the range of cognitive machinery that might support it. To most evolutionary biologists and behavioral ecologists, information transfer constitutes a form of teaching if it meets three

---

[1] We do not make a theoretical distinction between teaching and pedagogy. While the term "teaching" is generally used with nonhuman animals, and the term "pedagogy" reserved for humans, we do not find it useful to adopt this usage when it is an open question which cognitive processes support the various teaching phenomena observed across species and over development. Instead, we use these two terms loosely and interchangeably, aiming to highlight a number of more subtle distinctions necessary to characterize how different cognitive systems support teaching and pedagogical learning.

[2] These are not mutually exclusive accounts. It is possible, for example, that domain-general learning mechanisms exploit domain-specific perceptual schemas and/or motivational processes (see Heyes, 2012b,d). Similarly, evolved learning constraints can still exhibit experience-dependent plasticity, perhaps interacting with knowledge acquired via domain-general learning mechanisms. We highlight these two views as opposing extremes, acknowledging the nuanced positions between them.

criteria. To qualify as teaching, (1) the knowledgeable individual must modify its behavior in the presence of a naïve observer, (2) incurring a cost (or at least gaining no immediate benefit) from doing so, and (3) the observer must acquire knowledge or skills more efficiently than it otherwise would (Caro and Hauser, 1992)[3]. Phenomena that satisfy these criteria have been documented infrequently, but there are, nonetheless, examples in several distantly related nonhuman species (ants, *Temnothorax albipennis*: Franks and Richardson, 2006; meerkats, *Suricata suricatta*: Thornton and McAulliffe, 2006; pied-babblers, *Turdoides bicolor*: Raihani and Ridley, 2008). We argue that, despite these advances, further research is needed to understand the cognitive processes supporting teaching in these species and, in particular, the mechanisms by which naïve individuals learn from teaching when it occurs.

Although interest in teaching is shared across evolutionary and developmental disciplines, these lines of research have operated largely independently, yielding distinct emphases in each field. Studies of nonhuman animals tend to focus on whether and how a teacher caters its behavior to an audience to facilitate learning, whereas studies of human children focus on how the learner's assumptions allow the learner to effectively exploit pedagogical communication. In this review, we aim to integrate these perspectives by identifying cases of active information transfer in human and nonhuman animals and discussing what is known about the mechanisms that support these behaviors. We then suggest how our understanding of teaching in human ontogeny might take lessons from evolutionary theories, and from particular examples of teaching in nonhuman species, and propose ways in which findings and models from developmental psychology could inform our understanding of the evolution of social transmission mechanisms.

Learning from Others: Characterizing the Problem

All learners face the problem of generalizing from past observations or experiences to future, unseen events. Suppose, for example, that we observe that a new toy makes a novel sound. Given this observation, we must identify the actions or interventions needed to generate the sound and the set of objects that might share this sound-making property, as well as make numerous other inferences about whether the toy has additional hidden properties, and whether the sound is accidental, or central to the object's function. Recently, Bayesian models have been used as a flexible framework for specifying inductive inference problems at a computational level (Marr, 1982), formalizing how prior knowledge in a domain can interact with observed data to support such underdetermined inferences (Griffiths, Kemp, and Tenenbaum, 2008). While each of the inferences above is

---

[3] Notice that this definition does not provide criteria for differentiating teaching from many forms of communication that result in learning, including ones we might not readily identify as teaching in humans. We assume that there exists no hard and fast line between teaching and other such forms of communication, but that teaching generally communicates declarative or procedural information that generalizes to other contexts (e.g., "elephants have trunks"), as opposed to transiently relevant, episodic information (e.g., "look, there is an elephant in the room").

underspecified by a single observation, learners can make rich, rational inferences on the basis of limited experience if assumptions are made about the processes by which the observations have been sampled (Bonawitz et al., 2011; Gweon et al., 2010; Shafto, Goodman, and Frank, 2012). Consider, for example, observing that the novel toy makes the sound when on its side, but learning this either from observing the wind blow over the toy, from a person intentionally placing the toy on its side, or from a person saying “let me show you how this works” before placing the toy on its side. In the latter pedagogical context, a rational learner can make especially strong inferences based on assumptions about others’ communicative intentions and tendencies (e.g., that teachers tend to communicate complete, relevant, and generic information). Computational models of social learning have used a Bayesian framework to demonstrate that pedagogical instruction licenses a stronger set of inferences than mere observation of another agent’s behavior, so long as a learner assumes that a teacher is knowledgeable and provides representative, maximally informative data (Shafto and Goodman, 2008; Shafto, Goodman, Gerstle, and Ladusaw, 2010; Shafto et al., 2012).

Why Study Teaching in Children?

While these computational models show that a rational learner can generalize efficiently by making assumptions about pedagogical communication, this does not mean that all forms of teaching meet these assumptions, nor that making such assumptions is necessary for a learner to benefit from teaching. A teacher’s behavior might directly modulate attentional or motivational factors that facilitate learning, even if the learner has no prior assumptions to recruit when interpreting teachings acts. Do children have intuitive knowledge of pedagogical communication that constrains the inferences made from information provided in pedagogical contexts? Or does learning from human pedagogy rely on simpler attentional mechanisms that merely function to highlight relevant, generic information? Recent studies with human children and infants, reviewed in the following section, have aimed to distinguish between these possibilities.

These developmental studies also aim to shed light on the origins of receptivity to pedagogy. One possibility is that children’s ability to learn efficiently in social contexts, and from teaching specifically, is itself learned over the first few years of life. As children learn specific facts about people and their communicative tendencies, they might construct a “theory of pedagogy” that supports interpreting and learning from others’ pedagogical acts. The alternative proposed by Csibra, Gergely, and colleagues (the NP hypothesis; Csibra and Gergely, 2009, 2011) is that human infants have an innate receptivity to teaching that allows them to capitalize on the information provided in pedagogical contexts. This would involve genetically-adapted mechanisms for detecting acts of ostensive communication (via features like direct gaze and infant directed speech), and inferential constraints, already evident in infancy, that dictate the patterns of generalization made when these cues are present. Studies of humans across early development provide one method for examining the contributions of innate structure and experience to the emergence of effective pedagogical learning (see section titled “Relationship to Cognitive Architecture” for further discussion of the connection between innateness and specialized mechanisms).

Evidence for Receptivity to Teaching in Children and Infants

Do humans in fact learn from teaching by making assumptions about the nature of pedagogical communication? An emerging body of research suggests that children's learning is modulated in nuanced ways by the context in which data are obtained, and that children make uniquely strong inferences under pedagogical conditions. In an experiment by Bonawitz and colleagues (2011), for example, preschool children viewed an experimenter demonstrate one of several functional affordances on a novel toy (e.g., pulling a squeaker to make a sound) and were then given an opportunity to play themselves. In some conditions, the experimenter presented the information pedagogically (e.g., "This is my toy. I'm going to show you how my toy works"). In other conditions, the same affordance information was provided, but the demonstration violated key pedagogical assumptions, either because the toy was new to the experimenter, or because the experimenter was interrupted mid-demonstration. Children in the latter two conditions tended to explore and discover non-demonstrated functions of the toy. Children in the first condition, however, restricted exploration to the demonstrated function. These children apparently assumed that a teacher would communicate the full functionality of the toy, and thus inferred that the demonstrated affordance was its sole function.

Based on these and related findings (Buchsbaum, Griffiths, Gopnik, and Shafto, 2011; Butler and Markman, 2012; Gelman, Ware, Manczak, and Graham, 2013), it appears that children make assumptions about the communicative properties of teaching episodes (e.g., that a pedagogical demonstration will exhaust the functionality of a novel object). Further evidence that children (and adults) expect teachers to provide complete and representative information comes from the finding that learners evaluate an individual negatively for omitting relevant information when teaching others (Gweon, Pelton, and Schulz, 2011). Thus, by preschool years, children are able to take into account a range of relevant variables when interpreting communicative acts (i.e., whether the teacher completed the demonstration or was interrupted), and make flexible inferences about the relevance and scope of the knowledge provided in these contexts.

The presence of these abilities in human children is consistent with the view that humans possess an innate adaptation for receiving pedagogical communication, but also with the view that children construct the relevant assumptions (e.g., that teachers are knowledgeable and helpful, and communicate complete, relevant, generalizable information) over the first several years of life. Both views predict that children will learn efficiently from teachers, but make different predictions about the developmental trajectory of these pedagogy-specific inferences; the proposal that such inferences are learned implies an initial period in which infants, new to pedagogical contexts, will fail to make them, whereas the NP hypothesis predicts that such assumptions will be in place from the earliest exposure to pedagogical communication.

Attempts by proponents of the NP hypothesis to characterize infants' understanding of pedagogy focus on two key claims. First, adults tend to communicate with infants in systematic ways, and it is argued that infants are innately attuned to these behaviors, leading them to attend preferentially to instances of pedagogical communication (Csibra and Gergely, 2006). Consistent with this claim of communicative sensitivity, there is ample evidence that human infants attend to infant-directed speech (Pegg, Werker, and McLeod, 1992) and eye contact (Farroni, Csibra, Simion, and Johnson, 2002) and that these perceptual sensitivities play a non-trivial role in information transfer, eliciting gaze-following and joint attention from infants (Deligianni, Senju, Gergely, and Csibra, 2011; Senju and Csibra, 2008; Senju, Csibra, and Johnson, 2008).

The second, more controversial claim is that infants, like older children, represent the communicative intent behind these signals and make specific assumptions about the information being communicated, which guide subsequent processing and generalization of that information (Csibra, 2010; Csibra and Gergely, 2006, 2009, 2011). According to the NP hypothesis, the cues described above lead an infant to infer that information communicated will be shared across individuals, new or relevant to the learner himself, and generic or kind-generalizable (that is, will hold across instances of a type or category, rather than pertaining to the current exemplar alone). To support this proposal empirically, one must show not only that infants learn differently in pedagogical contexts, but also that they do so via specific assumptions about the information communicated through pedagogy. Evidence for this second claim rests on several key findings.

First, infants preferentially encode stable, category-relevant features of objects when engaged communicatively (Yoon, Johnson, and Csibra, 2008). In a violation-of-expectation paradigm (in which infants look longer at unexpected events), infants were repeatedly shown events in which a person either pointed to an object (communicative condition) or reached for it (non-communicative condition), followed by test events in which the object changed either location or identity. If infants expect that pedagogical communication will provide new knowledge that generalizes across contexts, ostensive communicative cues might lead infants to preferentially encode the identity of the referent, rather than transient properties that will not generalize to other instances (i.e., the referent's location). Consistent with these predictions, infants look longer when there is a location change following non-communicative reaching, but look longer when there is an identity change following communicative pointing. In a related study, Futo and colleagues find that infants map novel functions onto objects (evidenced in their use of function to individuate spatiotemporally ambiguous objects) only if the objects' affordances are demonstrated with the cues described above (Futo, Téglás, Csibra, and Gergely, 2010). These findings suggest that the information an infant encodes when observing a given object differs depending on whether the infant has been engaged pedagogically, or is simply observing the intentional action of another.

Further evidence for pedagogy-specific learning biases in infancy comes from a study by Topál and colleagues showing that ostensive communicative cues modulate perseveration in Piaget's A-not-B task (Topál, Gergely, Miklósi, Erdőhegyi, and Csibra, 2008). In this classic paradigm, an experimenter repeatedly hides a toy in a location and allows the child to search for it. After repeatedly retrieving the toy from the first location, infants continue to search that location even when they witness the experimenter hide the toy in a new place. While this perseverative error is normally explained in terms of the cognitive control required to inhibit returning to the previously successful search location, Topál and colleagues (2008) find that the effect depends heavily on the social-communicative context of the hiding event. When the experimenter exhibits ostensive cues while hiding the toy originally, infants make the perseverative error 81% of the time. When these communicative cues are absent, error rates drop by half. This effect of communicative context is interpreted as evidence that infants encode the ostensively signaled hiding event as conveying generic information (i.e., "toys are hidden here"), which continues to influence search behavior in subsequent trials[4].

Taken together, these findings suggest that infants encode different information in pedagogical and non-pedagogical contexts. But are the data described above clear evidence that infants represent the communicative intentions behind pedagogical overtures, and that their learning depends on specific assumptions about the nature of the information intentionally provided by others? Furthermore, the NP hypothesis asserts that humans learn from pedagogy via an adaptation that is unique to our species, a claim that obviously cannot be resolved with evidence from human infants alone. Experiments with nonhuman animals are necessary to assess whether forms of social learning present in early infancy are indeed fundamentally different from those observed in other species. In the following sections, we expand upon these developmental data and consider possible interpretations in light of findings with nonhuman animals. We begin by reviewing empirical phenomena from nonhuman animals, and discussing evolutionary theory relevant to teaching.

Evidence for Teaching in Nonhuman Animals

In evolutionary biology and behavioral ecology, there has been a long-standing interest in the forms of social transmission present across taxa. Recent research with several wild and captive animal populations has revealed behaviors that meet Caro and Hauser's (1992) functional definition, suggesting that teaching is a key form of social learning in some nonhuman systems. One example of teaching in a nonhuman species is the transmission of prey-handling skills in wild meerkats (Thornton and McAuliffe, 2006). Meerkat pups must learn to handle a range of mobile and potentially dangerous prey, and adult meerkats facilitate this learning process by bringing killed or disabled prey to young pups, and providing increasingly intact prey as pups mature. Critically, adults engage in these behaviors only when begging pups are present (criterion 1), and incur no obvious immediate benefit from provisioning pups in this way (criterion 2). This creates opportunities for pups to safely practice prey handling, and experimental interventions confirm that these behaviors facilitate skill acquisition (criterion 3).

The three functional criteria are also met by the case of tandem running in ants (*Temnothorax albipennis*). Knowledgeable individuals guide naïve nest-mates to food sources by running in tandem, slowing their pace as naïve individuals learn the route

---

[4] Similar effects have been found with domesticated dogs (*Canis familiaris*) (see Topál, Gergely, Erdhöhegyi, Csibra, and Miklósi, 2009), however only in human infants is the effect robust to change in the identity of the experimenter. This is taken as evidence that humans, but not dogs, interpret the communicative act as providing generic information that holds irrespective of the individuals involved.

(Franks and Richardson, 2006; Richardson, Sleeman, McNamara, Houston, and Franks, 2007). The leader modifies its behavior based on feedback from the follower, taking longer to obtain food than it would in the absence of a follower, and the naïve individual locates the food source more quickly than without a leader. Furthermore, the follower takes a more efficient path home, suggesting that tandem running not only facilitates the outbound journey, but results in learning the location of the food source with respect to the nest. Other examples include the teaching of food calls in pied-babblers (Raihani and Ridley, 2008), and putative evidence in several other taxa (elephants [*Loxodonta africana*], Bates et al., 2010; tamarins [*Saguinus oedipus*], Humle and Snowdon, 2008; see Byrne and Rapaport, 2011).

Importantly, these examples are classified as teaching according to a functional definition that makes no assumptions about the intentions or internal representations of the teacher or learner (Caro and Hauser, 1992). Operationalizing teaching in a mechanism-independent way has led to skepticism about the relevance of animal findings to an understanding of human teaching (Csibra, 2007). The value of a strict, outcome-based definition of teaching has been a topic of recent debate (Byrne and Rappaport, 2011; Rapaport and Byrne, 2012; see response from Thornton and McAuliffe, 2012. Here, we simply assume that the utility of a particular operationalization depends on what exactly is at stake in the research (for example, do we want to identify an animal model for human teaching, or to understand how social learning problems are solved in diverse neural architectures?). Because these research goals vary across disciplines and individual studies, and are only occasionally made explicit, we begin our discussion of how developmental and comparative approaches may benefit one another by reviewing a few central motivations of evolutionary biologists and comparative psychologists studying teaching.

Why Study Teaching in Animals?

One motivation for investigating teaching across non-human taxa is to uncover the selection pressures that lead to the evolution of different forms of social transmission, and the ecological or life-history conditions under which different behaviors manifest (Fogarty, Strimling, and Laland, 2011; Laland, Odling-Smee, and Gilbert, 2008; Leadbetter and Chittka, 2007; Thornton and Clutton-Brock, 2011; Thornton and Raihani, 2008). From an evolutionary dynamics perspective, active teaching presents several puzzles (Dawkins and Krebs, 1978). Although mechanisms for learning from others would serve a clear adaptive function, the advantage of mechanisms for actively transmitting knowledge to others are less obvious. Active teaching constitutes a form of cooperation (Hoppitt et al., 2008; Thornton and Raihani, 2008) and is therefore of interest to researchers exploring how seemingly altruistic behaviors can be evolutionarily stable strategies (Hamilton, 1964; Nowak, 2006; Trivers, 1971). Teaching might, for example, be more likely in eusocial or cooperatively breeding societies, due to kin selection (Burkhart, Hrdy, and van Schaik, 2009; Rapaport, 2006), which could partially explain the taxonomic distribution of documented teaching behaviors.

To characterize the evolution of teaching, we also need an account of the informational content that tends to require active transmission from a knowledgeable individual in the first place (see Fogarty et al., 2011 for a more comprehensive discussion). Each of the documented animal examples involves teaching of a distinct type of information: procedural skills in the case of meerkat prey-handling, a location in the case of tandem running ants, and a general association in the case of pied-babblers (for discussion of different forms of teaching, see Caro and Hauser, 1992). In the case of humans, Csibra and Gergely (2009, 2011) have argued for the role of active communication in transmitting functional or means-end information (relevant to exploiting and manufacturing causally opaque cultural artifacts like tools), and in spreading conventional knowledge systems (rituals, religions, food practices, languages, etc.). They argue that it was specifically the technological complexity of hominin artifact cultures that created selection pressure for rapid and robust transmission mechanisms. Similarly, Tomasello and colleagues have argued that intentional teaching is a key means by which our species sustains cumulative culture and cultural ratcheting (Tomasello, 1999; see also Dean, Kendal, Schapiro, Thierry, and Laland, 2012). However, studies with nonhuman animals suggest that the presence of causally opaque cultural products like tools and social conventions is not necessary for the evolution of all forms of teaching.

While work in evolutionary biology has traditionally been concerned with ultimate rather than proximate levels of analysis (Tinbergen, 1963), understanding the cognitive processes that underlie behavioral phenomena will be crucial to understanding their phylogenetic origins (Chittka, Rossiter, Skorupski, and Fernando, in 2012; MacLean et al., 2011). A key question for understanding the evolution of cognition concerns the specificity with which natural selection adapts systems to their environment (for a range of views, see Cosmides and Tooby, 1987; Heyes, 2012c; Hirschfield and Gelman, 1994). When does natural selection favor specific adaptations to a particular set of information-processing demands, and under what conditions are general-purpose resources exploited? In defense of a strongly domain-general view, Heyes has recruited comparative and neuroscientific evidence that social learning covaries with nonsocial learning across taxa, can be observed even in asocial species, and exhibits many of the key signatures of associative processes known to operate across disparate domains (Heyes, 2012b,d; see also Behrens, Hunt, Woolrich, and Rushworth, 2008). On the other hand, in meerkats and tandem-running ants, teaching presumably relies on mechanisms precisely catered to very specific learning problems of prey-handling and food localization in each species (Franks and Richardson, 2006; Thornton and McAulliffe, 2006). While it is unknown whether these mechanisms could, under the right environmental pressures, be co-opted to meet other transmission needs, the rigidity of these teaching phenomena (see below) suggests content-specific adaptations that are limited in scope. Identifying the conditions under which a more flexible multi-purpose teaching mechanism might evolve remains an important project for evolutionary biologists (see proposal in Csibra and Gergely, 2011).

Teaching also presents an interesting case for understanding the range of cognitive processes than could meet a particular functional demand, and the diversity of neuroarchitectures that would support these processes (Chittka et al., in press; Leadbetter and Chittka, 2007). In general, a contribution of comparative psychology can be to test hypotheses about which psychological mechanisms are necessary to support a behavior of interest (see, for example, Mendes, Rakoczy, and Call, 2008; Phillips and Santos, 2007 on the necessity of linguistic representations for kind-based object individuation). Based on the forms of teaching evident in humans, it has been suggested that effective teaching involves abstract inferential abilities, including representations of the knowledge and beliefs of the learner (Premack and Premack, 1996; Tomasello, 1999). Others focus on morphological prerequisites, suggesting that forms of social learning might be uniquely supported by the expanded primate neocortex (Humphrey, 1976; Reader and Laland, 2002). Evidence of teaching in insect and avian systems creates demand for more nuanced articulations of such claims. Because it is not always clear what sort of representations are prerequisites for a given behavior (see, for example, Cruse and Wehner, 2011), identifying manifestations of an ability across taxa can reveal the diversity of solutions to a basic functional or computational problem. Furthermore, when convergent evolution leads to analogous cognitive mechanisms across taxa, it may be useful to characterize how these are implemented in distinct neuroarchitectures (Emery and Clayton, 2004)[5].

Indeed, cases of teaching in nonhuman animals speak to the diversity of solutions to this class of learning problem. Consider, for example, the meerkat's ability to provide skill-appropriate prey in response to begging calls (Thornton and McAullife, 2006). A key question is whether meerkats possess an abstract representation of the pup's skill level, for which begging calls are just one reliable cue, or whether they respond to begging calls directly, without an intermediate inference about skill level. Empirical evidence falls in favor of the latter interpretation. When presented playbacks of begging calls from pups differing in age from those actually in the group, helpers provision based on the age to which recorded calls corresponded, failing to use other relevant information to provision the pups appropriately (for example, knowledge that the pup successfully handled intact prey previously). Such research characterizing how, in the abstract, an organism could solve a problem (along with an account the actual algorithms a neural system could implement to do so) will be crucial to understanding the evolution of functionally defined abilities like teaching (see Chittka et al., 2012; Thornton and Lukas, 2012).

Research on teaching can also shed light on the range of proximate mechanisms that have evolved to sustain cooperative strategies more broadly. Do prosocial or other-regarding motivations play a necessary role in teaching? Some claim that an intrinsic motivation to share information and mental states drives teaching in humans (Tomasello, Carpenter, Call, Behne, and Moll, 2005, Byrne and Rappaport, 2011) and perhaps other cooperatively breeding species (Burkhart and van Schaik, 2010). It is quite likely, however, that active information transfer occurs in the absence of empathic motivations (see relevant discussion in Vasconcelos, Hollis, Nowbahari, and Kacelnik, 2012). There is currently little data to speak to this question; whether any nonhuman species share information with a

---

[5] Of course, for certain purposes, homology is essential. The areas of human cognition that have been most fruitfully and rigorously characterized (i.e., the visual system) have clearly benefited from the neurophysiological methods afforded by the primate model (van Essen, Anderson, and Felleman, 1992). For this purpose, establishing common mechanisms is highly relevant, as the insights from the animal model to the human system will depend on the degree of homology between the neural substrates.

motivation to increase the naïve individual's wellbeing remains an open question for empirical research.

Human Pedagogy: Insights from Nonhuman Animals

Having laid out some key aims of research on the evolution of teaching, and some existing empirical findings with nonhuman animals, we now ask whether this body of work can constrain or inform the study of pedagogy in human learners. What can developmental psychologists learn from the study of teaching in other taxa? In interpreting experiments with human infants, the comparative perspective may be especially useful. As mentioned above, comparative evidence has been used to argue that many forms of social learning may depend on domain-general associative processes (Heyes, 2012b,d) rather than mechanisms evolved for the purpose of social transmission. This question remains unanswered with respect to our species' receptivity to pedagogy: Do humans have an innate adaptation for learning from teachers, triggered by a set of communicative signals for which evolution has prepared us? Or do humans rely on powerful domain-general learning mechanisms that enable them to quickly learn about human communication, recruiting previous teaching experiences to interpret subsequent communicative acts?

In considering the mechanisms supporting social learning across species, Heyes (2012b) emphasizes the distinction between perceptual mechanisms and learning or inference processes themselves, pointing out that general-purpose learning mechanisms can rely on input analyzers specialized for social information. That is, a species might have attentional or perceptual mechanisms (e.g., gaze-detectors) that facilitate the detection and exploitation of social information, while the subsequent processing of that information (e.g., categorization, inference, storage in long term memory) might depend on fully domain-general resources. However, on Heyes' view, even the input mechanisms themselves can be developed in a domain-general fashion, via experience with social stimuli reliably signaling fitness-relevant resources (see for example, Leadbetter and Chittka, 2007; Raihani and Ridley, 2008). As with other species, these questions remain wide-open with respect to human learning, and domain-general alternatives should be considered in parallel to the intriguing proposal put forth by Csibra and Gergely (2009).

Whereas Heyes (2012b) endorses a view in which social learning phenomena depend on resources that are both domain-general and representationally impoverished (i.e., associative mechanisms), it is also possible that organisms "learn to learn" socially via domain-general mechanisms that operate over abstract, structured representations (see Kemp, Perfors, and Tenenbaum, 2007). Another possibility is that input analyzers specific to ostensive communication might interact with mechanisms that are specific to the social domain, but broader in scope than the learning mechanisms assumed in the NP hypothesis (e.g., innate systems for representing goal-directed actions: Wood, Glynn, Phillips, and Hauser, 2007; social relationships: Seyfarth and Cheney, 2003). On the latter view, unique and powerful forms of social learning could emerge as newly evolved perceptual sensitivities combine in piece-meal fashion with phylogenetically older social-cognitive machinery.

Beyond the issues of domain-specificity and innateness, the presence of effective teaching in nonhuman animals highlights the question of whether the representations supporting receptivity to pedagogy in human infants need be as rich as those posited by the NP hypothesis. While the studies reviewed above suggest infants are attuned to communicative signals of others, and that ostensive-communicative cues facilitate generic knowledge acquisition, it is not clear that these phenomena reflect the sophisticated inferences that have been modeled and empirically demonstrated in older children. In the animal literature, there is considerable evidence that complex, adaptive behaviors can be supported by relatively lean mechanisms (Leadbetter and Chittka, 2007), and specifically that instances of effective teaching can depend on fairly simple stimulus-response mappings (Thornton and McAuliffe, 2012). For example, in the case of a meerkat adult modifying prey to match a pup's skill level, playback experiments indicate that the link between begging calls and prey modification is rigid and direct, likely relying on a mechanism that transmits very particular species-relevant information without high-level inferential abilities on the part of the teacher or recipient. With humans, as with other animals, we must distinguish between social-gating mechanisms (Kuhl, 2007) that might *function* to transmit generic knowledge from teacher to learner in pedagogical contexts, and mechanisms that do so by actually *representing* communicative intent and assumptions about the nature of the information provided.

The developmental experiments reviewed above appear to be open to multiple interpretations. The finding by Topal and colleagues (2009), for example, could be explained in terms of an attentional mechanism that leads to a stronger, more robust representation of the initial hiding event under pedagogical conditions. In the absence of pedagogical cues, working memory representations might be transient and weak, such that prior locations do not compete, and infants successfully search in the current location. When accompanied by pedagogical cues, both location representations are maintained, creating competition and a need to inhibit return to the prior location. Thus, the effect of social context might not depend on any sort of inference about the relevance or genericity of the initial hiding location, but simply a change in the strength of the location representation at the time of encoding.

As for Yoon and colleagues' (2008) finding that infants selectively encode kind-relevant information in a communicative context, it is possible that infants actually have a general tendency to encode and detect changes in the salient surface features of an object (evidence that features are not used for numerical individuation does not directly bear on this issue). On the other hand, the object's location becomes a highly relevant property when the target of a goal-directed reaching action. Thus, it is possible that detection of surface features in the communicative context reflects a baseline encoding of this featural information, with the detection of location information in a reaching context being an effect of the reaching cue specifically. Both interpretations are plausible, and a baseline condition is needed to distinguish between these possibilities[6].

Further experiments might seek to dissociate ostensive cues themselves from the

---

[6] Thanks to Elizabeth Spelke for pointing out this interpretation.

inference that the cues are supposed to license. What sort of information can infants exploit when evaluating the relevance of a given communicative event? As reviewed above, preschool-aged children distinguish between a communicative act that is successfully completed and an act that is interrupted (Bonawitz et al., 2011). Similarly, Bonawitz and colleagues find that children can learn from pedagogical overtures directed at a third-party who shares similar knowledge to the self (a child, in comparison to an adult), suggesting that direct eye-contact and contingency are not essential, but that children can flexibly evaluate the relevance of the information provided. Can infants identify the intention to teach in a similarly flexible, rational manner, even when the prototypical cues are lacking or misleading? Younger children (18-months) learn from third-party social interactions even when the target is an adult (Meltzoff and Brooks, 2008), so it remains to be seen whether toddlers could exploit other cues to the self-relevance of the information provided. Similarly, infants might discount ostensive cues if provided with evidence that these cues do not signal self-relevant information, for example, if given evidence that the communicative act was actually intended for a person with a different knowledge state than the infant. Such flexibility would provide stronger evidence for the view that infants represent the information as intended for them, and make inferences on this basis[7].

These experiments are designed to test the most extreme view, that infants form highly abstract representations of communicative intent, integrating a range of superficially disparate pieces of evidence to infer that pedagogical assumptions hold. However, a number of intermediate possibilities exist, and experiments exploring the scope and flexibility of these abilities will be essential. Here, we do not argue for any particular interpretation of the existing data, but simply wish to note the range of alternatives that are currently open. The animal literature attests to the power of natural selection to shape functional transmission behaviors without abstract inferential mechanisms (Leadbetter and Chittka, 2007, Thornton and McAuliffe, 2012), and the developmental community would do well to consider leaner mechanisms that could support human social learning as well[8].

The Evolution of Teaching: Insights from Human Development

Thus far, we have argued that the study of human pedagogy would benefit from the empirical phenomena and theoretical insights of recent work with nonhuman animals. But the evolutionary biologists among us hope that information can flow in the opposite direction as well. We now consider potential contributions of developmental psychology to

---

[7] Of course, a domain-specific learning mechanism of the kind proposed by Csibra and Gergely (2009) need not integrate all relevant information. Indeed, a key characteristic of domain-specific systems is their automaticity, and their reliance on a restricted set of inputs (Fodor, 1983). The question is whether infants' learning depends on a representation of communicative intent, as opposed to a more direct mapping between communicative cues and certain encoding strategies. Flexible integration of multiple cues to communicative relevance is just one way to argue for the former.

[8] See Heyes, 2012a on the role of parsimony in evaluating comparative data.

understanding the evolution of teaching and the mechanisms supporting teaching across taxa. First, although the distinctive feature of teaching is that the knowledgeable individual plays an active role in transmission, developmental studies highlight the fact that the term "active" is equally relevant for the learner. A complete account of teaching will involve characterizing the processes involved in instructing others and those involved in capitalizing on instruction received. Recent developmental and computational modeling work has emphasized the cues by which humans identify teaching contexts (Csibra and Gergely, 2006), and how the inferences they make in these contexts depend upon assumptions about the nature of information provided (Shafto and Goodman, 2008). These are potentially fruitful areas for future research with nonhuman animals.

Above, we reviewed evidence that human infants may be specifically attuned to behaviors that accompany communicative or pedagogical acts by adults. Do other species exhibit behavioral signatures that signal a mode of active information transfer, and if so, do naïve individuals exploit these cues to identify opportunities for learning? If teaching evolved in response to the substantial costs or dangers associated with individual exploration, selection may have favored rapid identification of contexts in which these dangers are ameliorated. Given that ostensive cues appear to have co-evolved with corresponding attentional biases in humans, it seems plausible that similar coordination would exist in other species. The meerkat pup, for example, would do well to identify prey that have been selected for them specifically (and thus are likely to be in a state manageable for their skill level) and to avoid unmodified prey, or prey modified for a more experienced pup. Consistent with this, Thornton and McAuliffe (2006) find that pups are more likely to interact with prey provisioned by a conspecific helper than identical prey provisioned by an experimenter, without a helper nearby.

It is possible that meerkats actually represent the fact that prey provisioned by conspecifics tend to be manageable for their skill level, and engage in otherwise risky exploratory behaviors because the teaching context suggests it is safe to do so. However, it is also possible that pups simply avoid food provided by a foreign individual. Manipulations analogous to those used by Bonawitz and colleagues (2011) might help to distinguish these possibilities. For example, Bonawitz found that children treat a pedagogical demonstration differently if the demonstration is interrupted. Are meerkats similarly sensitive to whether the provisioner was interrupted while disabling the prey? Children also selectively learn from teaching directed at a third party when the information is likely to be relevant to the self (i.e., when the demonstration is directed towards another child, but not when directed towards an adult). Could an experiment be designed to test whether meerkats also prefer to handle prey modified for a comparably skilled pup compared to prey modified for a more experienced individual?

Similar experiments could be designed with tandem running ants. Would ants explore less when taught the location of food (compared to a social foraging situation that did not involve teaching), analogous to the way human children engage in restricted exploration following what they assume to be a complete and representative demonstration (Bonawitz et al., 2011)? Restricted exploratory behavior seems to be a key signature of teaching in humans, and one that is intimately related to its primary function of restricting the learner's hypothesis space to support rapid learning (Bonawitz et al., 2011). Might this be a recurring phenomenon across taxa? Suppose an ant encounters a suboptimal food source either by following a tandem guide or by simply arriving at the source at the same time as another individual—would ants that discovered the food themselves be more likely to forgo low quality food and continue foraging, compared to those guided to this food patch? And if so, could experiments be designed to distinguish the effects of teaching from more general social influences on foraging? While Franks and colleagues have characterized some of the flexibility in ants' tandem running behavior (Franklin et al., 2011; Richardson et al., 2007), it is currently unknown whether ants would be sensitive to the various nuanced distinctions that modulate learning in human children. Probing the flexibility and sophistication of teaching across taxa will be crucial to understanding the evolution of different forms of social transmission, and the mechanisms supporting teaching in other species.

Recall also Gweon's studies in which children expect complete and representative information from teachers, and negatively evaluate others for omitting relevant information in third party interactions (Gweon et al., 2011). As discussed above, evolutionary theorists have often conceptualized teaching as a form of cooperation (Burkart and van Schaik, 2010; Thornton and Raihani, 2008). It appears that, in humans, the cooperative nature of teaching is actually represented by the recipient, and the expectation of helpful, representative information is part of what enables efficient transmission of knowledge in pedagogical contexts. Do other species anticipate the informational benefits of teaching interactions, and represent the interaction as a cooperative one? These questions may be crucial to understanding the evolution of teaching, as the learner's assumptions about the helpfulness of a teacher can make teaching a uniquely efficient mode of social transmission (Shafto et al., 2012).

Another key claim from the developmental literature is that teaching evolved to facilitate the transmission of cultural knowledge, and that children are specifically adapted to learn generic, conventional information from others (Cisbra and Gergely, 2011). Under this view, humans expect that the information provided during ostensive communication will not pertain only to the immediate present, but will generalize across people and time. Given that this generalizability is argued to be uniquely human, is there any evidence that nonhuman animals are biased to learn generic information from others? Meerkat pups clearly learn procedural knowledge that they generalize to future encounters with prey items, and it is possible that animals learn regularities that enable generalization beyond the particular teaching instances. For example, the ant could learn across tandem running instances that food sources tend to be located near certain environmental features, and this association could then facilitate future foraging.

However, these examples involve multiple instances of teaching, and the NP hypothesis is primarily concerned with the ability to generalize in the absence of exposure to regularities that would otherwise be necessary to warrant such inferences (Cisbra and Gergely, 2011). It is proposed that human learners can generalize from a single teaching instance (e.g., if taught a new fact about a dog, assume it to be true of all dogs, rather than idiosyncratic to the particular referent) based on an assumption that teachers provide generic, representative information (Csibra and Gergely, 2011). Bonawitz and colleagues (2011) provide evidence that preschool children do possess inductive constraints powerful enough to support rapid learning under pedagogical conditions (see also Futo et al., 2010; Gelman et al., 2012; Shafto et al., 2010). Future research with nonhuman animals might explore whether teaching in nonhuman animals ever supports this rapid sort of learning. Adapting the computational approaches that have been used to formalize learning and inference in children will be useful for assessing whether animal learning is best fit by models that assume pedagogical sampling, or those containing a weaker set of assumptions (see Shafto et al., 2012).

Relationship to Cognitive Architecture

While this review has focused primarily on questions regarding the *origins* of the ability to learn from teaching, the theories discussed herein can be associated with distinct claims about the underlying architecture as well. For example, if receptivity to pedagogy is constructed via domain-general learning mechanisms that detect regularities in the communicative interactions experienced over development, one might expect that the acquired knowledge will be represented and exploited using general-purpose cognitive machinery. Across domains, we interpret observations in light of our prior knowledge of the domain (e.g., interpret water droplets as rain rather than a sprinkler when the sky is gray, due to our intuitive knowledge of weather). One possibility is that the neural machinery that represents and exploits regularities in pedagogical communication to interpret subsequent communicative acts is highly general (it might operate across domains as diverse as weather prediction and pedagogical communication, or it might rely on mechanisms that are specific to social or mental state information [e.g., Saxe and Kanwisher, 2003], but not to pedagogical communication specifically). This perspective might also predict that these representations will be flexible, such that reliable communicative tendencies can be learned and unlearned based purely on the learners' experiences in relevant communicative contexts.

Nativist accounts, on the other hand, tend to be associated with claims of domain-specific systems operating with some rigidity throughout the lifespan (Spelke, 2003). Under this view, the abilities that emerge early in development are the result of functionally specialized neural circuits that were shaped by natural selection to solve evolutionarily relevant problems in efficient, content-specific ways. Thus, in the case of pedagogy, an innate adaptation might involve a dedicated neural mechanism for receiving communicative instruction from teachers. A strong version of this view might also predict that learners would respond to a fixed set of inputs (those that were communicatively relevant in the environment in which the system evolved), and do so in a relatively automatic, inflexible way. Rather than taking into account any and all relevant information to interpret diverse communicative acts, the cues exploited in pedagogical contexts might be limited and relatively indefeasible.

Do findings from developmental and evolutionary disciplines bear directly on these questions regarding the specificity of the neural architecture, and the flexibility of its operation? Many theorists have assumed that there will be reliable relationships between the origins of particular abilities (built from rich innate structure or from domain-general learning mechanisms), the neural localization of these abilities (specialized neural circuits, or recruitment of non-specific neural resources), and the signatures that characterize the relevant processing (fast, automatic, and constrained, or controlled, defeasible, and flexible) (Fodor, 1983). While this is an intriguing possibility, the relationship between these dimensions may not be so straightforward; learning can yield localized neural modules (e.g., VWFA in humans: see relevant discussion in Kanwisher, 2010), and innate content could, in principle, be represented and recruited for inferences in a way that is common across domains. Thus, we argue that future empirical and conceptual work is needed to clarify the relationship between these dimensions, and to establish the relevance of evolutionary and developmental data to claims regarding the underlying cognitive architecture.

Conclusions

The last decade has seen heightened interest in the development and evolution of teaching. However, extant dialogue between disciplines has generally expressed skepticism about the extent to which comparative data contributes to the study of human teaching (Byrne and Rapaport, 2011; Csibra, 2007). We welcome these concerns, and attribute them to ambiguity in the aims and assumptions underlying much current work in comparative psychology (see Boesch, 2007; Thornton and Lukas, 2012). It is increasingly common for researchers to compare performance of children and other animals, yet it is not always clear what is to be gained from these comparisons. In this review, we aimed to expand upon these critical analyses of the relationship between human and nonhuman teaching. Although the correspondence between teaching in humans and nonhuman animals is limited, we find the comparative approach to nonetheless be useful. Findings from these different disciplines can be mutually informative so long as researchers are explicit about what is assumed and at stake in the comparisons.

It is important to note that in arguing for an evolutionary perspective on human teaching, we are not endorsing a narrowly adaptationist approach. Rather we aim to delineate the numerous ways in which our evolved cognitive architecture might support active forms of social learning. We find it premature to favor one or another alternative based on the evidence available at this stage, but aimed to identify key distinctions and open questions, and suggest ways of distinguishing between these viable alternatives. Characterizing the evolution of the human mind will involve integrating theories and findings from evolutionary biology, comparative and developmental psychology, cognitive neuroscience, and computational modeling to understand how natural selection shapes neural systems to solve fitness-relevant problems in their environment.

We propose that continued research on teaching across phylogeny and ontogeny will be crucial to understanding the diverse forms social learning can take, and the possible mechanisms that can support effective transmission of different types of information. With respect to humans, it may be crucial to understanding our species-specific ability to maintain a rich, cumulative cultural inheritance (Fogarty et al., 2011; Tomasello, 1999). Humans inhabit dynamic, complex, and richly diverse environments, constructing intellectual, political, and cultural institutions that massively outstrip the achievements of even closely related species, and exhibit ratcheting over time. A central problem for evolutionary and developmental researchers is to characterize the key cognitive ingredients that drive these vast, undeniable differences (Heyes, 2012d; Tomasello et al., 2005). The study of teaching, along with other forms of social transmission, has the potential to yield progress on this question by characterizing how changes occurring over phylogenetic timescales support forms of learning that widen this gulf over ontogeny.

Acknowledgements: Thanks to the Origins of Cognition reading group for helpful discussion, and to Harvard's Mind, Brain, and Behavior initiative for funding the group. We wish to thank Alex Thornton, Hyowon Gweon, and Elizabeth Spelke for helpful discussion and comments on an earlier draft of the manuscript.

Received 15 August 2012; Revision submitted 24 October 2012; Accepted 26 October 2012

References

Bates, L. A., Handford, R., Lee, P. C., Njiraini, N., Poole, J. H., Sayialel, K., . . . Byrne, R. W. (2010). Why do African elephants (*Loxodonta africana*) simulate oestrus? An analysis of longitudinal data. *PLoS One, 5*, e10052.

Behrens, T., Hunt, L. T, Woolrich, M. W., and Rushworth, M. F. S (2008). Associative learning of social value. *Nature, 456*, 245-249.

Boesch, C. (2007). What makes us human (*Homo sapiens*)? The challenge of cognitive cross-species comparison. *Journal of Comparative Psychology, 121*, 227-240.

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., and Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition, 120,* 322-330.

Boyd, R., and Richardson, P. (1985). *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.

Buchsbaum, D., Griffiths, T. L., Gopnik, A., and Shafto, P. (2011). The influence of statistical and pedagogical cues on children's imitation of causal action sequences. *Cognition, 120*, 331-340.

Burkart, J. M., Hrdy, S. B., and van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology, 18*, 175-186.

Burkart, J. M., and Van Schaik, C. P. (2010). Cognitive consequences of cooperative breeding in primates. *Animal cognition, 13*, 1-19.

Butler, L. P., and Markman, E. M. (2012). Preschoolers use intentional and pedagogical cues to guide inductive inferences and exploration. *Child Development, 83*, 1416-1428.

Byrne, R. W., and Rapaport, L. G. (2011). What are we learning from teaching? *Animal Behaviour, 82*, 1207-1211.

Caro, T. M., and Hauser, M. D. (1992). Is there teaching in nonhuman animals? *Quarterly Review in Biology, 67*, 151-174.

Chittka, L., Rossiter, S. J., Skorupski, P., and Fernando, C. (2012). What is comparable in comparative cognition? *Philosophical Transactions of the Royal Society B, 367*, 2677-2685.

Cosmides, L., and Tooby, J. (1987). From evolution to behavior: Evolutionary psychology as the missing link. In J. Dupre (Ed.), *The latest on the best: Essays on evolution and optimality*. Cambridge, MA: The MIT Press.

Cruse, H., and Wehner, R. (2011). No need for a cognitive map: Decentralized memory for insect navigation. *PloS Computational Biology, 7,* e1002009.

Csibra, G. (2007). Teachers in the wild. *Trends in Cognitive Sciences, 11*, 95-96.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind and Language, 25*, 141-168.

Csibra, G., and Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In Y. Munakata and M. H. Johnson (Eds.), *Processes of change in brain and cognitive development* (pp. 249-274). New York: Oxford University Press.

Csibra, G., and Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*, 148-153.

Csibra, G., and Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B, 366*, 1149-1157.

Danchin, E., Giraldeau, L., Valone, T. J., and Wagner, R. H. (2004). Public information: From nosy neighbors to cultural evolution. *Science, 305*, 487-491.

Dawkins, R., and Krebs, J. R. (1978). Animal signals: Information or manipulation. In J. R. Krebs and N. B. Davies (Eds.), *Behavioural Ecology: An evolutionary approach*. Oxford: Blackwell Scientific Publications.

Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B., and Laland, K. N. (2012). Identification of the social and cognitive processes underlying human cumulative culture. *Science, 335*, 1114-1118.

Deligianni, F., Senju, A., Gergely, G., and Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-month-old infants. *Developmental Psychology, 47*, 1499-1503.

Emery, N. J., and Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in corvids and apes. *Science, 306*, 1903-1907.

Farroni, T., Csibra, G., Simion, F., and Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences USA, 99*, 9602-9605.

Fodor, J. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.

Fogarty, L., Strimling, P., and Laland, K. N. (2011). The evolution of teaching. *Evolution, 65*, 2760-2770.

Franklin, E. L., Richardson, T. O., Sendova-Franks, A. B., Robinson, E. J. H., and Franks, N. R. (2011). Blinkered teaching: Tandem running by visually impaired ants. *Behavioral Ecology and Sociobiology, 65*, 569-579.

Franks, N. R., and Richardson, T. (2006). Teaching in tandem-running ants. *Nature, 439*, 153.

Futó, J., Téglás, E., Csibra, G., and Gergely, G. (2010). Communicative function demonstration induces kind-based artifact representation in preverbal infants. *Cognition, 117*, 1-8.

Galef, B. G. (1976). A discussion of tradition and social learning in vertebrates. *Advances* *in the Study of Behavior, 6*, 77.

Galef, B. G. J., and Giraldeau, L. A. (2001). Social influences on foraging in vertebrates: Causal mechanisms and adaptive functions. *Animal Behaviour, 61*, 3-15.

Gelman, S. A., Ware, E. A., Manczak, E. M., and Graham, S. A. (2013). Children's sensitivity to the knowledge expressed in pedagogical and non-pedagogical contexts. *Developmental Psychology, 49*, 491-504.

Griffiths, T. L., Kemp, C., and Tenenbaum, J. B. (2008). Bayesian models of cognition. In R. Sun (Ed.), *Cambridge handbook of computational cognitive modeling*. New York: Cambridge University Press.

Gweon, H., Pelton, H., and Schulz, L. E. (2011). Adults and school-aged children accurately evaluate sins of omission in pedagogical contexts. In *Proceedings of the 33rd annual conference of the cognitive science society*, Boston, MA.

Gweon, H., Tenenbaum, J. B., and Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences USA, 107,* 9066-9071.

Hamilton, W. D. (1964). The evolution of social behavior. *Journal of Theoretical Biology, 7*, 1-52.

Heyes, C. M. (2012a). Simple minds: A qualified defence of associative learning. *Philosophical Transactions of the Royal Society B, 367*, 2695-2703.

Heyes, C. M. (2012b). What's social about social learning? *Journal of Comparative Psychology, 126*, 193-202.

Heyes, C. M. (2012c). New thinking about the evolution of human cognition. *Philosophical Transactions of the Royal Society B, 367*, 2091-2096.

Heyes, C. M. (2012d). Grist and mills: Cultural inheritance of cultural learning. *Philosophical Transactions of the Royal Society B, 367*, 2181-2191.

Hirschfield, L. A., and Gelman, S. A. (1994). *Mapping the mind: Domain specificity in cognition and culture*. New York: Cambridge University Press.

Hoppitt, W. J. E., Brown, G. R., Kendal, R., Rendell, L., Thornton, A., Webster, M. M., and Laland, K. (2008). Lessons from animal teaching. *Trends in Ecology and Evolution, 23*, 486-493.

Hoppitt, W. J. E., and Laland, K. N. (2008). Social processes influencing learning in animals: A review of the evidence. *Advances in the Study of Behavior, 38,*105-165

Humle, T., and Snowdon, C. T. (2008). Socially biased learning in the acquisition of a complex foraging task in juvenile cottontop tamarins, *Saguinus oedipus*. *Animal Behaviour, 75*, 267-277.

Humphrey, N. K. (1976). The social function of intellect. In P. P. G. Bateson and R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303-317). New York: Cambridge University Press.

Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences, 107*, 11163-11170.

Kemp, C., Perfors, A., and Tenenbaum, J. B. (2007). Learning over hypotheses with hierarchical Bayesian models. *Developmental Science, 10*, 307-321.

Kuhl, P. K. (2007). Is speech learning "gated" by the social brain? *Developmental Science,* *10*, 110-120.

Laland, K. N., Atton, N., and Webster, M. M. (2011). From fish to fashion: Experimental and theoretical insights into the evolution of culture. *Philosophical Transactions of the Royal Society B*, 366, 958-968.

Laland, K. N., Odling-Smee, F. J., and Gilbert, S. F. (2008). EvoDevo and niche construction: Building bridges. *Journal of Experimental Zoology B, 310*, 549-566.

Leadbeater, E., and Chittka, L. (2007). Social learning in insects: From miniature brains to consensus building. *Current biology, 17*(R), 703-713.

MacLean, E., Matthews, L. J., Hare, B. A., Nunn, C. L., Anderson, R. C., Aureli, F., . . . Wobber, V. (2011). How does cognition evolve? Phylogenetic comparative psychology. *Animal Cognition, 15*, 223-238.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Meltzoff, A. N., and Brooks, R. (2008). Self-experience as a mechanism for learning about others: A training study in social cognition. *Developmental Psychology*, *44*, 1257-1265.

Mendes, N., Rakoczy, H., and Call, J. (2008). Ape metaphysics: Object individuation without language. *Cognition*, *106*, 730-749.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science, 314*, 1560-1563.

Pegg, J. E., Werker, J. E., and McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week old infants. *Infant Behavior and Development*, *15*, 325-345.

Phillips, W., and Santos, L. (2007). Evidence of kind representations in the absence of language: Experiments with rhesus monkeys (*Macaca mulatta*). *Cognition, 102*, 455-463.

Premack, D., and Premack, A. J. (1996). Why animals lack pedagogy and some cultures have more of it than others. In D. R. Olson and N. Torrance (Eds.), *Handbook of education and human development: New models of learning, teaching and schooling* (pp. 302-323). Oxford: Blackwell.

Raihani, N. J., and Ridley, A. R. (2008). Experimental evidence for teaching in wild pied babblers. *Animal Behavior, 75*, 3-11.

Rapaport, L. G. (2006). Parenting behaviour: Babbling bird teachers? *Current Biology*, *16*, 675-677.

Rapaport, L. G., and Byrne, R. W. (2012). Reply to Thornton and McAuliffe (2012). *Animal Behavior*, http://dx.doi.org/10.1016/j.anbehav.2012.06.013.

Reader, S., and Laland, K. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences USA, 99*, 4436-4441.

Rendell, L., Fogarty, L., Hoppitt, W. J. E., Morgan, T. J. H., Webster, M. M., and Laland, K. N. (2011). Cognitive culture: Theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*, *15*, 68-76.

Richardson, T. O., Sleeman, P. A., McNamara, J. M., Houston, A. I., and Franks, N. R. (2007). Teaching with evaluation in ants. *Current Biology, 17*, 1520-1526.

Saxe, R., and Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in “theory of mind”. *NeuroImage*, *19*, 1835-1842.

Senju, A., and Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology, 18*, 668-671.

Senju, A., Csibra, G., and Johnson, M. H. (2008). Understanding the referential nature of looking: Infants' preference for object-directed gaze. *Cognition, 108*, 303-319.

Seyfarth, R. M., and Cheney, D. L. (2003). Signalers and receivers in animal communication. *Annual Review of Psychology, 54*, 145-173.

Shafto, P., and Goodman, N. D. (July, 2008). *Teaching games: Statistical sampling assumptions for pedagogical situations*. In Proceedings of the 30th annual conference of the cognitive science society, Washington, DC.

Shafto, P., Goodman, N. D., and Frank, M. C. (2012). Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science, 7*, 341-351.

Shafto, P., Goodman, N. D., Gerstle, B., and Ladusaw, F. (August, 2010). *Prior expectations in pedagogical situations*. In Proceedings of the 32nd annual conference of the Cognitive Science Society, Portland, OR.

Slagsvold, T., and Wiebe, K. L. (2011). Social learning in birds and its role in shaping a foraging niche. *Philosophical Transactions of the Royal Society B, 366*, 969-977.

Spelke, E. S. (2003). Core knowledge. In N. Kanwisher and J. Duncan (Eds.), *Attention and performance, Vol. 20: Functional neuroimaging of visual cognition*. Cambridge, MA: MIT Press.

Strauss, S., Ziv, M., and Stein, A. (2002). Teaching as a natural cognition and its relations to preschoolers' developing theory of mind. *Cognitive Development, 17*, 1473-1487.

Thornton, A., and Clutton-Brock, T. (2011). Social learning and the development of individual and group behaviour in mammal societies. *Philosophical transactions of the Royal Society B, 366*, 978-987.

Thornton, A., and Lukas, D. (2012). Individual variation in cognitive performance: Developmental and evolutionary perspectives. *Philosophical Transactions of the Royal Society B, 367*, 2773-2783.

Thornton, A., and McAuliffe, K. (2006). Teaching in wild meerkats. *Science, 313*, 227-229.

Thornton, A., and McAuliffe, K. (2012). Teaching can teach us a lot. *Animal Behaviour, 83*, e6-e9.

Thornton, A., and Raihani, N. J. (2008). The evolution of teaching. *Animal Behaviour, 75*, 1823-1836.

Tinbergen, N. (1963). On aims and methods in Ethology, *Zeitschrift für Tierpsychologie, 20*, 410-433.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*, 675-735.

Topál, J., Gergely, G., Erdhöhegyi, A., Csibra, G., and Miklósi, A. (2009). Differential sensitivity to human communication in dogs, wolves, and human infants. *Science, 325*, 1269-1272.

Topál, J., Gergely, G., Miklósi, Á., Erdőhegyi, Á., and Csibra, G. (2008) Infants' perseverative search errors are induced by pragmatic misinterpretation. *Science, 321*, 1831-1834.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology, 46*, 35-57.

Van Essen, D. C., Anderson, C. H., and Felleman, D. J. (1992). Information processing in the primate visual system: An integrated systems perspective. *Science, 255*, 419-423.

Vasconcelos, M., Hollis, K., Nowbahari, E., and Kacelnik, A. (2012). Prosociality without empathy. *Biology Letters, 8,* 910-912.

Whiten, A. (2011). The scope of culture in chimpanzees, humans and ancestral apes. *Philosophical Transactions of the Royal Society B, 366*, 997-1007.

Wood, J. N., Glynn, D. D., Phillips, B. C., and Hauser, M. D. (2007). The perception of rational, goal-directed action in nonhuman primates. *Science, 317,* 1402-1405.

Yoon, J. M. D., Johnson, M. H., and Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences USA, 105*, 13690-13695.