5 Relevance-Based Emulation as a Prerequisite for Technical Innovation

György Gergely and Ildikó Király

Introduction

To understand the cumulative nature of human culture and its transmission across generations, it is important to shed light on the evolutionary origins and multiple determinants of humans' unique capacity to produce new, efficient, and innovative technological solutions to relevant problems. As a result of their increased efficiency, new innovative techniques are more likely to be selected, to be transmitted, and to stabilize in a culture, thereby extending the scope of the existing repertoire of cultural skills. The emergence of new and innovative technologies, therefore, contributes significantly to the cumulative nature of human cultural knowledge. The increasing scope and complexity of the repertoire of technological skills that are maintained and transmitted to new generations imply two different challenges that evolved mechanisms for transmitting cultural knowledge must solve. On the one hand, the transmission process must lead to the production of sufficient variability among the reproduced variants of the original technological skills if it is to enable the emergence of new and more efficient innovative technological solutions. On the other hand, the transmission mechanisms must also be able to reproduce cultural skills sufficiently faithfully for these to be successfully transmitted and maintained in a cumulative cultural repertoire.

The adaptive solution satisfying these two requirements may lie in the special design properties of evolved cultural transmission mechanisms. In order to account for transmission processes that produce both variability in alternative form variants and learning strategies generating rigid and conservative motor replicas of acquired technical skills, recent theoretical accounts postulate a strict teleological dichotomy between instrumental actions and ritual-like action kinds. In contrast, in this chapter we propose and argue for the existence of a *relevance-based emulative learning mechanism* by focusing on the central role played by *ostensive communication* in technical knowledge transmission. Our theoretical account differs in significant respects from the dichotomy-based approach because it focuses on the central role played by pedagogical ostensive communicative demonstrations—demonstrations that alternative accounts largely overlook.

Our approach is built on the cognitive foundations of human ostensive communication (Csibra and Gergely 2009, 2011; Heintz and Scott-Phillips 2022; Sperber and Wilson 2002). Our main claim is that when someone is communicatively addressed by knowledgeable

others, it induces an expectation of relevance in naive learners, which is sufficient to account for both flexible (and potentially innovative) as well as rigid, high-fidelity aspects of cultural transmission. Communicative demonstrations function to bring into focus those parts of an action that are manifested as relevant by ostensive communication and, as such, are highlighted to be reenacted and learned. Notably, ostensive communicative cues identify the context and conditions in which an action sequence is adequate for use.

For instance, imagine someone waving their hand in the air. This behavior could be interpreted differently in particular contexts. If one discovers that a fireplace is close to the person, one might think that the goal of the behavior is to avert the smoke, in which case this behavior would qualify as a transparent instrumental action. However, if one observes the same action—waving the hand—without detecting any justificatory physical aspect in the situation, then the goal of the action (e.g., possibly serving as a gesture to greet a social partner or to express respect) remains teleologically opaque to the juvenile. In this example, the same instrumental action could be interpreted and represented as being transparent or opaque simply as a function of the specific context in which it is performed.

Given that the physical-causal relations between the observed or demonstrated actions and their subsequent outcomes often appear opaque to naive learners, relying on trustworthy and knowledgeable social partners' relevance-guided communicative demonstrations is an efficient strategy that novices can exploit to figure out what is worth acquiring as relevant in particular contexts. In other words, when the goal and goal structure of a behavior are opaque, the ostensive demonstrative behavioral cues accompanying the performance of the opaque action can be used to inform the naive learner that, despite its apparent teleological opacity, the ostensive highlighting itself means that the action is relevant for the novice to acquire and reenact in the kind of contexts where it was demonstrated.

Cognitive Mechanisms for Knowledge Transmission

Humans are certainly unique among other social and cultural species in their remarkable ability to create, maintain, and transmit across generations an ever-growing body of knowledge. This knowledge includes instrumental skills, know-hows, innovative techniques, and technical inventions, as well as traditional, conventional, normative, and rule-governed forms of social activities. These various kinds of culturally transmitted action forms allow humans to pursue a variety of instrumental and social goals that serve different adaptive functions in their cultural community. Achieving even a partial understanding of the origins, evolution, and multiple determinants of the processes involved in human cumulative culture (e.g., generating innovations and successfully transmitting and stabilizing them across generations) is a huge scientific enterprise requiring a massively multidisciplinary approach and interdisciplinary cooperation.

The classical view shared by psychologists, cultural anthropologists, ethnographers, and (until recently) archaeologists held that the obvious adaptive mechanism selected for the human transmission of novel cultural skills is the capacity for behavioral imitation, which allows for the faithful motor reproduction of novel action forms that naive human learners observe their culturally knowledgeable social partners performing (Tennie, Call, and Tomasello 2009; Legare et al. 2015; Boyd and Richerson 1996).

Similarly, many ethnographers and cultural anthropologists have claimed that novel cultural skills are learned in a straightforward manner by young children through passive participant observation, imitative copying, and practicing the imitatively reproduced behavioral forms with peers during joint play (Paradise and Rogoff 2009; Lancy, Gaskins, and Bock 2009; Lancy 1996). They argue that parental investment in the costly forms of active teaching, communicative demonstration, ostensive attention guidance to relevant properties, correction, feedback, or explanation are not cultural universals but culture-specific inventions related to the emergence of formal schooling (Lancy 1996) and practiced only in so-called WEIRD (Western, educated, industrialized, rich, and democratic) societies (Henrich, Heine, and Norenzayan 2010; for a contrary view, see Hewlett and Roulette 2016). Thus, the standard and still widespread view holds that humans' special adaptation for imitative learning by behavioral copying serves as the central and unique psychological mechanism of human cultural transmission, enabling us to reproduce and acquire novel skills.

This view has also received support from intriguing research on the striking phenomenon called "over-imitation" (Lyons, Young, and Keil 2007) in children from about four years onward and that remains present even in adults (Hoehl et al. 2019). The numerous studies on over-imitation appear to provide evidence that naive human cultural learners evolved a strong and spontaneous inclination to reenact faithful behavioral replicas of novel action sequences performed by adult models. In these experiments, children are presented with a transparent box and provided demonstrations of a sequence of novel actions used to obtain a goal object from the box, such as a sticker. These experiments intentionally vary an essential feature of human behavior: its causal transparency.[1] These studies revealed that even when the initial actions in a sequence are obviously causally unnecessary or irrelevant to the task (such as performing "magical" circular movements with a feather without making contact with the transparent box) and only the final action in the sequence is causally necessary and instrumentally relevant, children still readily and faithfully reproduce the whole sequence. Such a spontaneous inclination for over-imitation has been demonstrated by cross-cultural researchers in WEIRD societies and in a variety of non-WEIRD societies as well (Nielsen and Tomaselli 2010; Berl and Hewlett 2015). Even more surprising is that, upon being subsequently questioned about which of the imitated actions were necessary to retrieve the goal object and which were "silly," the children readily identified the causally irrelevant and unnecessary actions as the "silly" ones that they, nevertheless, faithfully reenacted (Lyons, Young, and Keil 2007).[2]

In sum, a variety of studies of imitative learning of novel actions converge in diagnosing humans' unique cognitive adaptation as a specialized behavior-copying imitation mechanism that naive cultural learners rely on to reproduce faithful motor replicas of novel and often causally opaque actions performed by knowledgeable cultural models. By enabling the faithful reproduction of actions and skills, this uniquely human mechanism has been, therefore, assumed to play a crucial role in the intergenerational cumulative growth and maintenance of a shared repertoire of skills within human social groups (Tomasello 1999).

At the same time, however, a growing body of evidence from developmental and comparative studies of imitative learning in humans and primates has led to the increasing recognition that the processes of imitative reenactment are often characterized by a significant degree of flexibility and variability. Indeed, while reproducing the same goal outcome as the observed novel action, learners frequently introduce alternative (and often less costly) action

variants of varying efficiency. This difference was first explicitly recognized by Michael Tomasello (1996, 1998), who proposed a distinction between two types of reenactment mechanisms for reproducing novel action skills during cultural transmission (Gergely and Csibra 2006).

Tomasello (1998) termed the *goal-* or *outcome-emulation mechanism* the more flexible kind of reenactment mechanism that can generate variable forms of actions reproducing the same goal outcome as the originally observed action. He situated this mechanism in contrast to the more rigid *imitative behavior-copying mechanism* that produces faithful motor replicas of observed means actions. In goal emulation, the learner's attention is focused primarily on reproducing the goal outcome of the modeled behavior, while the specific means actions performed by the model can be flexibly substituted by alternative action variants that realize the same goal state. In fact, some of these variants can prove to be more efficient than the original observed action and be retained in the social group. It was therefore hypothesized that the relative flexibility of processes of emulation and the consequent degree of variability in modified action forms may play an important role in the generation of innovative and increasingly efficient action routines—thereby cumulatively and adaptively enriching the cultural repertoires transmitted to future generations (Legare and Nielsen 2015; Legare et al. 2015).

Determinants and Functions of Rigidity versus Flexibility of Cultural Transmission Mechanisms

A theoretical proposal (Clegg and Legare 2016; Legare and Nielsen 2020) postulates two basic dichotomies that human cultural learners are assumed to possess. First, it is proposed that humans are cognitively adapted to distinguish and recognize two ontologically distinct types of actions: (a) *causally transparent*, efficient goal-directed actions serving instrumental functions (such as tool use), and (b) *causally opaque* conventional or traditional types of actions that are typically jointly and publicly performed practices serving social rather than instrumental functions (such as rituals or traditional joint activities; see Legare and Nielsen 2020).

The second dichotomy distinguishes between two alternative learning mechanisms that human cultural learners are hypothesized to possess (Clegg and Legare 2016; Legare and Nielsen 2020). The first is a *more flexible emulative action reproduction mechanism* that generates more variability and modifications in the reproduced action variants. This mechanism is hypothesized to be selectively triggered by the learner's recognition of the causal transparency of a novel goal-directed action that serves instrumental functions. The second kind of imitative learning capacity is the *more rigid behavior-copying mechanism that produces faithful motor replicas* of an observed novel action. This more conservative transmission mechanism is assumed to be selectively induced by the learner's recognition of the causal opacity of conventional or traditional noninstrumental action forms. This more rigid imitative motor action copying mechanism would generate significantly less variability in the novel cultural action forms that it produces (Legare 2019; Legare and Nielsen 2015; Legare et al. 2015; Tennie and van Schaik 2020).

This theoretical account attempts to explain the differential distributions of rigidity versus flexibility in the reproduction processes by linking the different kinds of actions (causally

transparent and opaque actions) to the two different types of social learning mechanisms (emulation and imitation). The emulation-based reenactment mechanism activated by the causal transparency of novel instrumental actions is hypothesized to generate more variability and alternative functional variants to solve the same goal, thereby supporting the emergence of "innovations."[3] These potentially more adaptive and novel variants can then be retained and stabilized within the population, fueling cumulative change in the action repertoire transmitted to subsequent generations (Clegg and Legare 2016; Legare and Nielsen 2015). In contrast, the apparent causal opacity of conventional or traditional action forms that are novel to the learner induces the more rigid imitative action copying mechanism, resulting in faithful motor action replicas with little variability. This conservative strategy is hypothesized to play an adaptive role by inhibiting the generation of disruptive variability of causally opaque action forms, thereby promoting successful cultural transmission and the stable maintenance of repertoires of conventional traditions. The suggestion here is that if such causally and teleologically opaque ritual acts induced the more flexible and emulative strategy, which would generate variability in the ritual behavior, the conventional behavior's recognition later would be endangered, thereby reducing the likelihood of the conventional action's successful maintenance and stability across generations (Kapitány and Nielsen 2019).

One of the central problems faced by this dichotomizing approach is that it largely overlooks the important fact that the repertoire of novel cultural actions to which naive social learners are exposed includes many goal-directed actions that serve primarily instrumental functions but nevertheless have constituent parts (lower-level actions or subgoals) that appear causally opaque to the learners (e.g., by violating the causal efficiency requirement of goal-directed instrumental actions). Equally importantly, partially causally opaque instrumental actions often serve traditional or ritualistic functions as well. For example, while eating food with one's hands is a highly efficient and causally transparent instrumental act, adults in various cultural traditions present the novice with less efficient normative manners and styles of consuming food (e.g., eating with knives and forks or using chopsticks). These alternative culture-specific variants of instrumental actions are more complex and contain causally and teleologically opaque parts, which include performing more costly and less instrumentally efficient actions than the alternative of eating food with one's hands does. Yet, their primary function is still obviously instrumental, while also serving social and traditional display functions as well.

Such transitive actions with "mixed" functions involve culturally normative, subefficient manners of action execution that are not causally transparent to the learner and are in fact more costly to perform than other equally available, more efficient, and often more familiar alternatives. The specific causal and functional properties of these actions must thus appear causally opaque for the juvenile cultural learner.[4] However, the dichotomy-based approach holds that observing the apparent causal opacity of a novel action form induces in the naive cultural learner a high-fidelity, rigid copying mechanism that produces exact motor replicas of the opaque target action, thereby generating little or no variability in the reproduced action forms. This prediction holds whether the causal opacity is detected in the action serving noninstrumental social functions or at the level of the causally opaque subactions or subgoals of a clearly goal-directed, larger action that serves both an instrumental and a social function.

Relevance-Based Emulation as the Mechanism of Cultural Learning

We propose an alternative view according to which relevance-based goal emulation serves as the dedicated adaptation for social learning, selected for acquiring and transmitting the repertoire of culturally shared action skills. This mechanism is induced by ostensive communicative manifestations of culturally relevant intentional actions irrespective of their degree of causal transparency or opacity.

According to our view, children are adapted to recognize the communicative gestures of knowledgeable adults as indicating pedagogical (or demonstrative) contexts in which new and relevant cultural information is made manifest. Ostensive communicative signals induce their addressee to segment the demonstrated action into constituent parts such as its component means actions, the subgoals that these means actions bring about, the final goal to which the means actions ultimately lead, and the specific context in which the action is performed overall. The pedagogical signals highlight for the naive learner the novel and relevant aspects of the segmented action sequence that should be learned and reproduced in a given context. This pedagogical guidance allows the learner to identify both the overall goal and, when highlighted, the relevant subgoals leading to it—whether or not these behaviors are instrumental. In what follows, we will discuss the available empirical evidence as well as new experimental data supporting the selective and inferential nature of the relevance-guided emulation mechanism.

Our studies focus on the transmission of a goal-directed instrumental action that involves a novel, causally "opaque," and subefficient subgoal. These studies demonstrate the fast-learning, long-term retention, and flexible production of alternative action variants of the causally opaque and subefficient elements of novel instrumental actions. We take these results as evidence in support of our claim that the dedicated psychological mechanism serving cultural transmission in humans is one of *selective, inferential, and relevance-guided emulation*. This mechanism involves the ability to *flexibly choose alternative means actions* to reproduce relevant subgoals manifested by the pedagogical demonstration of knowledgeable partners. In this sense, we argue that the transmission of novel, ostensively demonstrated goal-directed actions supports the production of alternative behaviors to achieve goals manifested as relevant, thus providing the basis of the capacity for innovative use of accumulated knowledge.

As such, our alternative proposal equally supports the claim that a flexible cultural (emulative) learning mechanism may provide variation in the reproduced actions, fueling the emergence and selection of more efficient versions of the original action. As these more efficient action variants stabilize within a population, they can lead to the cumulative enrichment of cultural action repertoires. Therefore, variability of reproduction is, indeed, an important source of the ratchet effect in cultural transmission, contributing to the cumulative nature of human culture that characterizes its transformations through cycles of intergenerational transmission (Boyd and Richerson 1996; Tomasello 1999).

Context-Sensitive Selective Imitation of Causally Opaque Instrumental Actions

In this section, we present some studies that aim to test the above hypotheses by using new versions of the "head touch paradigm," originally designed to investigate the imitative

reenactment of novel and causally opaque actions by infants. This paradigm (first presented in Gergely, Bekkering, and Király 2002) demonstrates that imitation is a selective, inferential, and context-sensitive learning mechanism.

In the original study, 14-month-old infants watched an adult sitting in front of a table with a touch-sensitive lamp on it. The experimenter first placed her hands on the table next to each side of the lamp and then performed an unusual and subefficient means action to illuminate the lamp: she bent over the lamp to press its touch-sensitive surface with her forehead (the "hands free" condition). A separate group of infants were tested in an alternative context condition where the model first pretended to be freezing, telling the infant that she was really cold, and so she put a blanket around her shoulders and held onto it tightly with both hands (the "hands occupied" condition). She then went on to demonstrate the very same unusual and subefficient head touch action to light up the lamp, bending over it and activating it with her forehead.

In the test phase, the infants were given the touch-sensitive lamp and were encouraged to play with it on their own. In the hands-free condition, most of them (69%) used their head to activate the lamp (cf. Meltzoff 1988). However, in the hands-occupied condition, only a small proportion of infants (21%) performed the head action to light the lamp; most of them just used the more efficient (but undemonstrated) method of pressing the light box with their free hands.

These patterns of selective imitation of the demonstrated head touch action indicated that infants were sensitive to the context in which the model presented the unusual action. In the hands-free condition, the model's subefficient and unusual head touch action must have appeared causally opaque to the infants, given that the demonstrator's hands were resting freely on the table and thus could have been used to press the lamp (a more efficient and familiar alternative means action to light up a lamp). Nevertheless, the demonstrator opted not to use her free hands, instead presenting to the infants the causally opaque (subefficient and more costly) head touch action. In contrast, in the hands-occupied condition, where the model's hands were not free (being occupied with holding the blanket around her shoulders), the demonstrator's head touch action must have appeared causally transparent to the infants, since using the head to activate the lamp was contextually justified as a causally efficient alternative action, given the constraints of her hands being occupied.

These findings have been interpreted as evidence that imitative reenactment is a selective, inferential, context-sensitive learning mechanism that relies on evaluating the relative efficiency of the target action observed as a function of the situational constraints on possible actions (cf. Gergely and Csibra 2003). In other words, when deciding whether or not to reenact a causally opaque behavioral component or subgoal of the novel action, the naive learner took into account the action's relative efficiency in obtaining the specific goal outcome as a function of the constraints imposed by the particular context (Gergely, Bekkering, and Király 2002; Gergely and Jacob 2012; Király, Csibra, and Gergely 2013).

Advocates of the dichotomy-based approach to cultural transmission (discussed above) could argue that it can account for these findings of selective reenactment insofar as it appears to support the central assumption of their theory—namely, that the causal opacity of an observed action induces high-fidelity, rigid behavior-copying in the cultural learner (Clegg and Legare 2016; Legare and Nielsen 2015). Additionally, these results appear to be in line with the idea that the faithful copying of novel instrumental actions is the result of a copy-when-uncertain strategy in social learning (Rendell et al. 2011; Toelch, Bach, and Dolan

2014). It has also been proposed that high-fidelity imitation may be so useful when learning novel but opaque behaviors that its benefits outweigh potential efficiency costs in the transmission process (McGuigan et al. 2007).

In contrast in the hands-occupied context, the demonstrated head touch action must have appeared *causally transparent* to the infants. Yet, despite its apparent causal transparency, the head touch action failed to be imitated. Instead, infants—whose own hands were unoccupied—chose far more frequently to use their free hands to produce the more efficient "hand touch" action to activate the lamp. Note that this finding appears hard to accommodate within the framework of the dichotomy-based approach (Legare et al. 2015; Legare and Nielsen 2015).

The challenges that need to be handled by a dedicated adaptive cultural learning mechanism are twofold. First, such mechanisms should support the fast-learning, long-term retention, and functionally adequate delayed reenactments of novel means actions. These aspects are the main signs of the reliable and long-lasting acquisition of a new action. Second, it should at the same time allow for the flexible and functionally appropriate generalization and selective reproduction of the newly acquired motor skill across a variety of relevant and novel contexts. Learning a new skill is only adequate if it is applicable and useful in future situations as well, and this requires the identification of new contexts where it is relevant.

These criteria represent a challenge for the dichotomy-based approach. In response to the first challenge, this view does not offer a means to distinguish instrumental transitive actions that contain *causally opaque* subcomponents while serving a primarily transparent instrumental function (while also serving social functions) from conventional and traditional action routines, which also involve causally opaque actions as well as actions allegedly serving only social and affiliative functions. Thus, the dichotomy-based approach faces a challenge in providing the criteria necessary for the adequate and flexible selection of adaptive means.

With regard to the second challenge, the cultural learning model proposed by the dichotomy-based approach does not adequately support generalization across functionally relevant new contexts. The dichotomy-based view proposes that based on its apparent causal opacity, naive cultural learners recognize and categorize the observed novel action as belonging to the domain of conventional, traditional actions that serve noninstrumental social functions. Because of their causal opacity, such actions induce and are acquired by conservative imitative behavior-copying, and their rigidly reproduced motor replicas are probably stored separately as conventional, traditional actions serving primarily social functions. As a result, instrumental transitive actions that contain causally opaque components could become miscategorized as actions serving only social and conventional functions. However, in order to provide advantageous solutions for similar problems in other instrumental contexts, the acquired novel action routines should be applicable in functionally relevant new situations as well. Yet it is unclear how the strategy of rigid imitative copying of causally opaque action skills could promote generalization across such contexts.

Our alternative approach, however, holds that there are in fact no distinct imitative mechanisms to support the acquisition of conventional as opposed to instrumental actions. Rather, the same learning mechanism is used to learn both opaque and transparent novel actions. In the case of the causally opaque behavioral components that are ostensively manifested by the demonstrator as culturally relevant, more elements of these causally opaque but relevant action components are likely to be reproduced, allowing better identification of appropriate contexts and, consequently, flexible generalization of behavior.

Let us investigate what strategies infants pursue when they attend closely to ostensively modeled behavioral elements in relation to the relevant context within which they are demonstrated. We propose that ostensive attention guidance to the relevant (though possibly opaque) action components highlighted by the communicative action manifestation, together with the experience of alternative variants and their repeated application in variable contexts, should provide the relevant informational basis to identify instrumental actions as separate from conventional ones in such contexts.

The Illusion of Imitation: Is There Imitative Form-Copying?

With the aim of investigating more closely the underlying mechanism behind the selective learning of causally opaque instrumental means actions, we (Király, Csibra, and Gergely 2013) ran follow-up studies using the head touch paradigm. They found that 14-month-olds selectively reenacted the novel, apparently arbitrary, and subefficient means action (lighting the touch lamp by contacting it with one's forehead) in the hands-free context condition only when the subefficient and opaque "head touch" action was demonstrated by an adult model addressing them in an ostensive communicative manner. This ostensive manifestation was interpreted by the naive learner as reflecting the adult's pedagogical intention (Csibra and Gergely 2011) to communicate that the causally subefficient and opaque means action was nevertheless culturally relevant and, as such, should be acquired by the naive learner. However, when the same action was performed but observed from a third-person perspective without being accompanied by ostensive cues of communication by the adult, infants did not reenact the causally opaque subefficient head touch action. Instead, they achieved the same goal more efficiently by using their hand to operate the touch-sensitive light box.

Király, Csibra, and Gergely (2013) proposed that the selective reenactment of the novel behavior observed in the communicative context—specifically, the imitation of the novel and arbitrary head touch means action in the hands-free context—demonstrates that young children are prepared to reenact and acquire novel actions even when their subefficient execution appears causally opaque. In the absence of the possibility of exploiting their individual learning strategies, they rely on the communicative signals of experienced others. Infants interpret the ostensive action demonstrations as pedagogically intended communicative manifestations of novel and (in spite of their causal subefficiency) culturally relevant means actions to be acquired. Because of the ostensive communicative signals that accompany the action demonstration, infants construe it as conveying new and relevant information that the demonstrator intends to communicate and not as a purely instrumental action (Csibra and Gergely 2011; Gergely and Jacob 2012; Altinok, Király, and Gergely 2022). The infants would thus interpret the situation as a teaching context where the demonstrated action manifested to them with ostensive communicative cues is intended to guide them in learning a relevant (if causally opaque) novel instrumental action.

In the hands-occupied condition, the obvious constraint of having the hands busy with another goal is sufficient for young children to form a coherent interpretation of the demonstrator's choice to use her head to light up the lamp. However, this is not the case in the hands-free condition. Instead, in this context, the child observer must search for an alternative explanation to understand the use of the head action. The communicative demonstration context induces in them the presumption of relevance (Sperber and Wilson 2002), leading

them to interpret the opaque and subefficient means action manifested as relevant (from some unspecified point of view) and, as such, a behavior to be acquired because it is culturally significant. As a result, in addition to interpreting the final instrumental goal as "lighting up the lamp," infants will also construe as relevant the apparently causally opaque subgoal (i.e., to achieve the final goal by making contact between the lamp and their forehead). This approach thus proposes that the ostensive communicative demonstration context can enrich the encoding of the overall goal-directed action by signaling that the specific manner of performing the means action (contacting the lamp with the head) is, in spite of its apparent subefficiency, the culturally relevant way of attaining the final goal. But is this interpretive process served by a rigid imitative copying of the demonstrated, causally opaque action?

A closer look at the performance of children allows us to answer this question. According to the dichotomy-based approach, when children turn to behavior-copying after observing a new, arbitrary, and causally opaque action used to attain a goal, their strategy would be to merely copy (or imitate) the performed behavior. With respect to the head touch action, this strategy implies that children would construe the head action as a causally opaque but successful way of bringing about an effect and would not use any other means to achieve the goal.

To pursue this question further, we first analyzed the reenactment behavior of children in relation to the goal object. We found that those children who performed the head action not only performed this novel means action, but they also performed the simpler (but undemonstrated) hand action to attain the goal—and did so without exception. In fact, the hand action in all cases preceded the head action and was successful in bringing about the effect (Király, Csibra, and Gergely 2013). This means that children encoded the overall goal of the situation; more importantly, they were inclined to try out alternative means that turned out to be more effective in attaining the same outcome.

Second, we coded the specific forms of head touch actions with the aim of assessing the potential variability in the particular manner of reenacting the modeled action. The level of variability could reflect the degree of fidelity of imitative behavioral responses and as such could be used as an indicator of the underlying mechanism of social learning. The results showed that children performed the head action with high variability. They used different parts of their head to contact (or approach) the lamp, including their cheeks, faces, mouth, eyes, and even their ears. Thirty-six percent of imitators performed two or even three different forms of head actions during the testing session, and only 11 percent produced a faithful version of touching the lamp with the forehead—however, they did so while also using other parts of their head to light up the lamp, contacting it with an ear or with the face (see figure 5.1). In addition, in some cases (25%), they only approached the lamp and never made the contact between their head and the lamp (Király, Csibra, and Gergely 2013).

Furthermore, children also alternated the way that they managed to bring the lamp into contact with or close to their head: they either followed the demonstrated version, leaning toward the lamp without moving it, or most interestingly, they grabbed the lamp by hand and lifted it up to touch their head. This intriguing variant of an emulative response was at first only a chance observation (because with overuse, the adhesive putty fixing the lamp to the table became loose, allowing the infants to grab and raise the lamp by hand). Based on this observation, we developed a further study (Chen, Király, and Gergely 2012).

Figure 5.1
The different ways of using the head to act on the lamp. *Source*: Photo by Ildikó Király.

Following the demonstration with a lamp fixed to the table, for the test phase different (though similar-looking) touch lamps were put on the table in front of the children, each being easy to lift. In this version, we could directly observe the variability that children revealed in choosing their specific version of means action during reenactment. We found that after the ostensive demonstration of the target act of leaning forward and contacting the lamp with the forehead, 85 percent of the imitators (i.e., 58% of the participants) lifted the lamp to their head. Moreover, in half of these cases, children performed both the lifting and the leaning-forward variants of the head action (Chen, Király, and Gergely 2012).

From the perspective of the dichotomy-based view, opaque action demonstrations should trigger imitative form-copying. According to this account, in the head touch paradigm, the arbitrary and opaque application of the head being bent down to touch the lamp should specifically and solely induce high-fidelity copying of the action sequence demonstrated. However, as described above, we found a great deal of variability, which allows us to suggest that children interpret the demonstration as a communicative action in which the demonstrator informs them about the instrumental goal and also about the subgoal relevant to achieving the final goal within the context (i.e., to activate the lamp by bending forward and touching it with the forehead).

These findings clarify the role of inferential processes involved in action analysis and reveal the important role of ostensive communication in enabling infants to represent the goal structure—that is, the overall goal and the specific means as the relevant subgoal—of novel actions even when the causal relations between the means and the end-states are causally opaque. We propose that the presumption of relevance induced by the ostensive cueing guides infants' interpretations of the relevant subgoal/final goal structure of the demonstrated action sequence. Indeed, it seems that while they reproduce the ostensively interpreted relevant goal structure of the manifested action sequence, infants continue to monitor the efficiency of actions when choosing, changing, or disregarding certain action elements, so long as these modifications leave constant—or make more efficient—the realization of the relevant subgoal/final goal structure. We further suggest that after the observation of a novel goal-directed action, infants can identify the goal of the action that they encoded. When they

are invited to reenact the action, they recall the encoded goal and (re)enact an action variant to attain the same goal. Most importantly, they encode the novel means as a subgoal when it is signaled as novel and relevant by the ostensive communicative context.

We take these results as evidence against the behavior-copying hypothesis of the dichotomy-based approach and argue instead that relevance-based emulation serves as a central mechanism driving reenactment. In this process, the subgoal is accentuated as relevant by the ostensive communication, keeping open the possibility that either (a) the subgoal is causally linked to the overall goal and serves as a placeholder for potential technical variation (which remains opaque to the infant), or (b) the subgoal is a social goal representing a socially accepted alternative manner of attaining the final goal.

In essence, we posit that there is only one form of cultural learning that infants use to learn both causally opaque and transparent behaviors: the emulation of encoded goals or goal structures. Yet, the richness and detailed nature of this encoding of goals is modulated and guided ostensively during its communicative demonstration. The presumption of relevance induced by the communicative situation highlights novel information in relation to the overall goal, but it does not necessarily disambiguate initially the exact sense in which the ostensively highlighted aspect of the action is novel and relevant. Indeed, further communicative exchanges allow for disambiguating in what sense the behavior is relevant (e.g., understanding whether the subgoal has instrumental merits or is rather a convention serving social purposes). This later disambiguation, however, requires some grounding points or bases for further elaboration. How then do children differentiate instrumental from conventional behavior during observation?

It is widely accepted that behavioral reenactment contributes to the transmission of conventions and can serve social functions itself (Legare et al. 2015; Over and Carpenter 2012, 2013; Watson-Jones and Legare 2016; Wen, Herrmann, and Legare 2016). From a young age, children attend to a variety of social and contextual cues to determine the goal of behaviors (Buchsbaum et al. 2011; Carpenter, Call, and Tomasello 2005). From this angle, the learning situation should highlight those contextual features that allow children to map when conventional manners are demonstrated and when there is more space for refinement of an instrumental action.

It is also a possibility that ostensive communication helps children to differentiate instrumental goals from conventional ones. In the following section, we present novel studies that directly investigate whether the communicative demonstration of a novel action can provide disambiguating cues by differentiating the relevant action contexts. We argue that these cues can help infants infer whether and how to reenact demonstrated novel actions and guide them toward evaluating the conventional versus instrumental functions served by the same action.

Further Studies to Provide Evidence in Support of Relevance-Driven Goal Emulation

We argue that in order to learn novel actions, infants need to be able to identify the relevant target actions. Moreover, for this purpose, infants need to rely on the active inferential guidance provided by the demonstrator's ostensive communicative gestures and manner of

manifesting the action (cf. Gergely 2007; Gergely and Csibra 2006; Csibra and Gergely 2011; Király, Csibra, and Gergely 2013). In the following experiments, we introduce modifications to previous experimental setups—namely, the "hands up" and "balls" conditions (Paulus et al. 2011)—in order to test whether the presumption of relevance induced by ostensive communication guides infants to infer and identify the relevant information that is applicable and generalizable for later use.

The Role of the Demonstrated Relevant Context in Interpreting the Manifested Opaque Means Action

Our first objective in this series of studies was to test the proposal that observing a causally opaque action leads to imitative behavior-copying (possibly as a result of the induced motor resonance of the infant's corresponding motor programs; see Paulus et al. 2011). Our second aim was to test the alternative hypothesis of our inference-based selective emulation model, which holds that ostensive demonstration of an opaque means action is interpreted by infants through context-sensitive inferences constrained and informed by the relevant aspects of the context in which the action takes place.

In proposing our alternative relevance-based emulation model, we predicted that the ostensive manifestation of the opaque head touch action would induce infants to attend to the relevant action context in which the head touch action is demonstrated. As a result, infants would interpret the demonstrated opaque head touch action of the model by evaluating it in relation to the relevant action context. In particular, we predicted that by varying relevant aspects of the demonstration context in which the very same opaque head touch action is observed, we could induce differential and selective reenactments of the opaque action by the infants. To achieve these aims, we borrowed and modified the experimental conditions that were initially designed to provide evidence for the motor resonance-based, automatic behavior-copying theory of imitative learning (see Paulus et al. 2011).

Study 1. Imitation of the Opaque Head Touch Action in the "Palms in Air" Demonstration Context

Markus Paulus and colleagues (2011) criticized the inference-based selective imitation account of the original "head touch" study (Gergely, Bekkering, and Király 2002) by suggesting that the reason infants failed to imitate the demonstrated "head action" in the "hands occupied" condition could be that 14-month-olds simply cannot bend over to touch the lamp with their head without supporting their body by putting their hands on the table. In fact, Paulus and colleagues suggested that infants imitated the "head touch" action in the "hands free" condition precisely because they observed that the experimenter put her hands on the table to support her body by leaning on them when bending forward to touch the light box with her head—that is, precisely the way that the infants were themselves constrained to perform the action. According to Paulus and colleagues, observing this configuration of the model's actions induced motor resonance in the infants' corresponding body parts and activated the motor imitation of both the model's body supporting hand actions and the bending over to touch the lamp with their forehead. So, contrary to the account provided by the inference-based rational imitation theory, infants did not infer that the model's hands were free. Instead, they observed that the hands were occupied by supporting the model's body during the head action.

Our version of the "hands up" condition described by Paulus and colleagues (2011) is the "palms in air" demonstration. In our version, the context differed from Gergely, Bekkering, and Király's (2002) original "hands free" condition. While in both conditions the model's hands were free when the head action was performed, in the "palms in air" condition, hands were not placed on the table—thus, they could not have been used to support the model's body while bending forward. After sitting down in front of the table and the touch lamp, the model in the "palms in air" condition demonstrated two different salient actions separately in a sequence. The model first presented a hand action extending her two hands toward the light box on the table while turning her palms upward midair (see figure 5.2). This hand action corresponds to the kind of semi-conventionalized ostensive referential manual gesture that humans often use to "show" or "highlight" an object or event as relevant for another to attend to. In everyday communication, this demonstrative manual gesture is often accompanied by some verbal referential expression, such as "Here!" in English or "Voilá!" in French—something that we have adopted in our ostensive communicative demonstration condition.

After the model's hands had finished their referential gesture highlighting the lamp (accompanied by the referential vocal gesture "Voilá!"), a slight pause followed, and then the model proceeded to perform a second action with her head. She bent forward from the waist and lit up the lamp by touching it with her forehead. During the performance of the head touch action, her hands remained stationary in their previous position (held with palms up midair). Therefore, the "palms in air" demonstration context provided clear temporal and contextual segmentation cues to help infants interpret the model's demonstrated hand gesture as a separate referential action that established a relevant context for interpreting the subsequent head touch means action. (Note also that the hands' referential act was clearly completed, and so the hands in the air in front of the infants' eyes were clearly "free" to be used for a new action.)

We presented two groups of 14-month-olds with this action sequence in two demonstration conditions: in a second-person ostensive communicative context and in a third-person noncommunicative observation context. In the communicative second-person condition, apart from providing temporal action segmentation cues, the demonstrator also addressed the infant through ostensive referential gestures and presented the action demonstrations in an ostensive way (i.e., in a slightly exaggerated "motionese" manner). This provided infants with ostensive highlighting and temporal parsing cues to guide them to separately interpret the initial hand action demonstration as an ostensive referential manual gesture. We hypothesized that the presence of these ostensively provided informative cues—similar to Gergely, Bekkering, and Király's (2002) original hands-free condition—would direct the infant to parse and interpret the hand gesture as forming part of the relevant action demonstration context rather than being part of the demonstrated head touch target action itself. In contrast, in the third-person noncommunicative observation context of the "palms in air" study, the demonstrator presented the exact same action sequence to a different group of 14-month-olds without any ostensive communicative gestures.

For the "palms in air" condition, our relevance-guided inferential account predicts that, guided by the provided ostensive signals and temporal parsing cues, infants will be able to infer the new and relevant information manifested for them by the ostensive demonstration

of the unusual head touch means action, and they will thus be able to learn and reenact it. For the third-person condition, our account predicts that without the presence of such ostensive communicative cues, there will be no imitative reenactment of the causally opaque and teleologically subefficient head touch action. In contrast, according to the dichotomy-based view, which claims that observing a causally opaque cultural action induces faithful behavior-copying, both conditions should induce a reenactment by the infant since both the communicative (second-person) and the noncommunicative (third-person) observation conditions present the exact same (partially opaque) action sequences to the infants.

Method

Participants. Twenty-six 14-month-old infants were recruited (two were excluded because of parental interference or fussiness). Twenty-four children were assigned to one of the two experimental conditions (12–12).

Test phase. The modeling phase in both conditions (the second-person communicative demonstration context and the third-person noncommunicative observation context) was immediately followed by the test phase, in which the infants received the light box. Infants were given 60 seconds to explore and play with the lamp.

Data Analysis and Scoring

The video records of the test phase were scored by two independent observers who were uninformed about which of the conditions the participants belonged to. The dependent measure was whether the infant attempted to perform the head-on-box action within a 60-second time window. An attempt was defined as either touching the lamp with the head or approaching the lamp with the head (e.g., leaning forward) within 10 centimeters or less (see Meltzoff 1988). We also coded for the direction of approach of the target action. The potential ways of approaching the lamp were either leaning forward or lifting up the lamp. The two coders' evaluations of the participants' performances were in 97 percent agreement (kappa = 0.94).

Results

The proportion of infants who performed the target action is presented in figure 5.2.

We compared the performance in the two conditions. The frequency of target action reenactment was lower in the third-person "noncommunicative observation context" condition than it was in the "communicative demonstration context" condition (Fisher exact $p = 0.05$). Calculating the odd ratio confirmed (OR = 5.431) that infants in the "communicative context" condition were more likely to reenact the head action in comparison to the group of infants in the "noncommunicative context" condition. As in previous studies (Gergely, Bekkering, and Király 2002; Paulus et al. 2011), at least one hand action preceded the head action in 92 percent of cases. The frequency of hand actions was 7.8 for one head touch within the first 60 seconds.

Interestingly, the head touches appeared in different forms than the modeled behavior. Most importantly, 30 percent of infants who performed the head action (three infants in the communicative demonstration context and one infant in the noncommunicative observation context) lifted the lamp up to their heads instead of leaning forward to touch it. Moreover, in 30 percent of cases, there was no contact between the approaching head and lamp (two

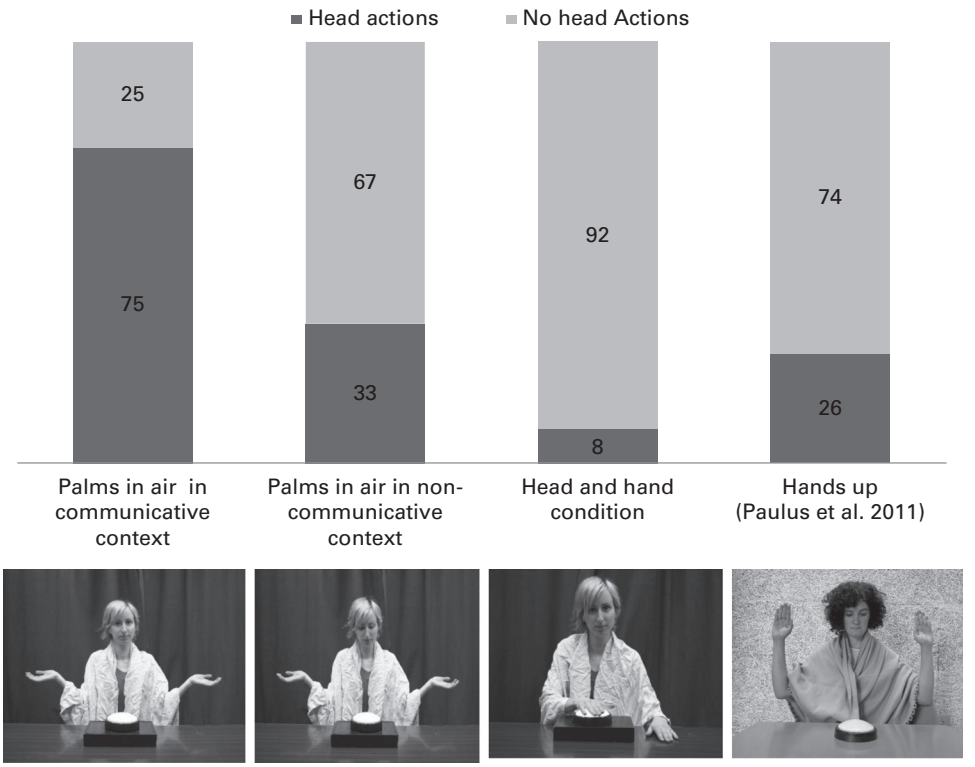


Figure 5.2
The proportion of imitators in each condition in study 1 and study 3. *Source*: Photo by Ildikó Király.

infants lifted the lamp up to the head but did not make physical contact with it, while two other infants bent forward but did not make contact with the lamp; all of these were in the communicative demonstration context).

These results confirm that infants performed voluntarily chosen variations of the originally observed behavior, rather than reenacting a matching motor replica of the observed action. The selective pattern of results in the different demonstration conditions provides further empirical evidence for the natural pedagogy account of learning (Gergely and Csibra 2006)—namely, that the presence of a demonstrator's ostensive and referential communicative signals addressing the infant is a critical factor that is necessary to induce the imitative reenactment of a novel and apparently subefficient means action.

Study 2. Situating the Goal in Context: The "Balls" Study

The main objective of this study was to demonstrate that ostension plays a crucial role in linking the overall goal of an action with its manifested subgoals, and that this integration is driven by the learner relying on the relevant aspects of the action context being manifested. Furthermore, we also aimed to show that the ostensive demonstration of opaque means actions of an instrumental transitive act could induce variability in the action, thereby leading to alternative actions when reproducing the opaque subgoal. We highlight the role of osten-

sive communicative and temporal parsing cues that could guide infants' interpretive inferences to identify the relevant aspects of the new information manifested to them given the demonstrated action context.

In the "balls" condition of Paulus and colleagues (2011), there were two softballs lying on the table next to the lamp. In the demonstration context preceding the manifestation of the opaque means action, the experimenter took a seat and played with the two softballs for approximately eight seconds. Then, keeping one ball in each hand, the experimenter put her hands on the table next to the lamp. From then on, the procedure followed exactly that of the hands-free condition in Gergely, Bekkering, and Király (2002), with the only difference being that the experimenter was holding the two softballs in her hands on the table while performing the opaque head action to activate the lamp.

In this condition, one can also argue that observing the hand action itself (i.e., putting the hands with balls on the table) may not be sufficient for the infant to infer if the hands are free or occupied. Such an inference must rely on and is constrained by the relevant aspects of the context in which the hand actions were demonstrated. Here, the model was playing with the balls and then stopped—a context in which infants could infer that her hands were now free to act (they do not necessarily have to continue holding the balls).

Nevertheless, to clearly disambiguate the interpretation of the relevant action and help the infants parse the manifested action sequence, we introduced two different versions of the demonstration context. In the "hands free resting on balls" condition, we followed the procedure of the "balls" condition of Paulus et al. (2011), except that in our study (1) the two balls were lying on two little plates on the table, and (2) after the model had put her hands with the balls next to the light box, she lifted her hands up without the balls for two to three seconds. The balls remained in the plates and could not roll away. Then the model put her hands down again, grasping the balls on the plates as before. After this short event, the model performed the head action with her hands resting on the balls. In this context, it was made explicit during the demonstration that the hands were free to act because they were not occupied with holding the balls so that they would not roll away.

In the "hands occupied with holding balls" condition, there were no plates next to the light box, and the model performed the exact same action sequence performed in the "hands free resting on balls" condition. Accordingly, when the model lifted her hands for a moment, the balls started to roll away, so she had to quickly reach back and grasp them again to keep them from moving farther. This situation unambiguously manifested that the hands were occupied and were not free to engage in another action. In both situations, however, the model's hands (with the balls in them) were placed on the table, so they could provide support for her body when she bent forward to touch the lamp with her head.

According to the motor resonance theory and the dichotomy-based view, there should be no difference in the number of imitators in the two conditions. However, the different situational constraints demonstrated relevant contextual information for the infant to infer whether the hands were free or occupied during the performance of the head touch action. This allowed the infants to interpret the head action as an efficient means to perform in the condition in which the hands were occupied with holding the ball ("hands occupied with holding balls" condition). In contrast, given the relevant contextual information demonstrating that the hands were free and could have been used to touch the lamp ("hands free resting

on balls" condition), infants could infer the demonstrated relevance manifested by performing the subefficient and causally opaque head touch action. This generates the prediction that the number of imitators should differ in the two conditions.

Method

Participants. Thirty 14-month-old infants were recruited; three of them were excluded from the final sample because of fussiness (n = 1), technical error (n = 1), and parental interference (n = 1). Participants were randomly assigned to the two experimental conditions. As a result, 14 infants were tested in the "hands free resting balls" condition, and 13 infants were tested in the "hands occupied with holding balls" condition.

Test phase. The test phase immediately followed the modeling phase in both conditions. The model pushed the lamp across the table in front of the infant and said, "It is your turn now! You can try it!" She encouraged the infant to play with it and stayed in the room. Infants were given 60 seconds to explore and play with the lamp.

Data Analysis and Scoring

The video records of the test phase were scored by two independent observers who were uninformed about which of the conditions the participants belonged to. The dependent measure was whether the infant attempted to perform the head action within a 60-second time window (as in study 1). The two coders' evaluations of the participants' performance was in 92 percent agreement (kappa = 0.85).

Results

The number and proportion of infants who performed the target action are presented in figure 5.3.

When we compared the performance in the two conditions, it was revealed that the frequency of target action tended to be lower in the "hands occupied with holding balls" condition than it was in the "hands free resting on balls" condition (Fisher exact p = 0.054). Odd ratio (OR = 5.177) examination revealed that the probability of performing a head

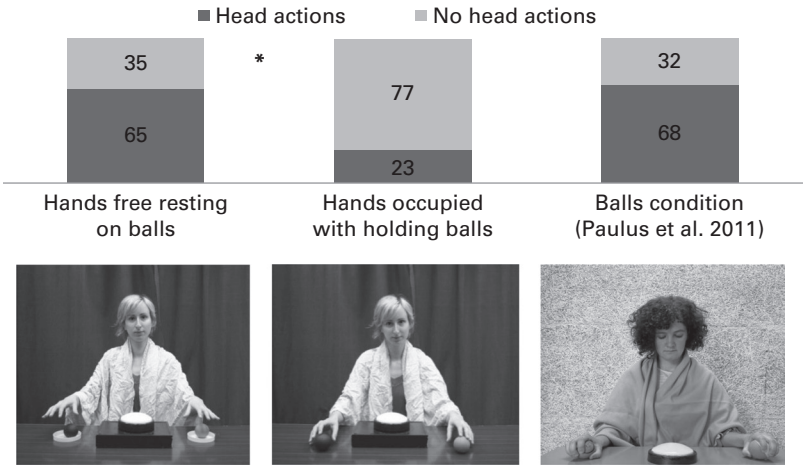


Figure 5.3
The proportion of imitators in each condition in study 2. *Source*: Photo by Ildikó Király.

touch is more likely in the "hands free resting on balls" than it is in the "hands occupied with holding balls" condition.

Hand touch actions preceded head touch action in 94 percent of cases. The frequency of hand touch actions was almost six for a head touch. Moreover, like in study 1, the head touch actions did not follow the modeled head touch with high fidelity. Intriguingly, in 50 percent of cases, infants lifted the lamp up to their heads instead of leaning forward to it (four infants in the "hands free resting on balls" condition and two infants in the "hands occupied with holding balls" performed the head touch action this way). Also, in 58 percent of cases (out of which 25%, or three infants, were lifters), there was no contact between the approaching head and the lamp.

Study 3. Control Condition: Ostensive Demonstration of Both the Head Touch and the Hand Touch Actions

Our main suggestion is that when children observe in a communicative context the ostensive demonstration of a novel subefficient goal-directed means action that cannot be justified as efficient, they will interpret it as a causally opaque subgoal that represents a socially relevant manner in which to attain the final goal in the demonstrated context. In the present experiment, we explore the further assumption that when children have already acquired an efficient (causally transparent) means action to operate an artifact (which is therefore not novel to them), they will not learn to reenact an alternative nonefficient (and so casually opaque) means action to achieve the same goal, even if it is presented in an ostensive communicative context (Pinkham and Jaswal 2011), since infants would infer from the ostensive demonstration that the hand action is a socially accepted and efficient means to attain the goal. However, here, the dichotomy-based approach would predict that children will imitate the head touch action, since in this view, seeing a causally opaque action automatically triggers this form of social learning.

Method

Participants. Fourteen 14-month-old infants were tested in the "head and hand" control condition.

Apparatus. The apparatus used was identical to that of study 1.

Modeling phase. Here the model demonstrated two actions (in ostensive-communicative context), while she also demonstrated that her hands were free (lying on the table next to the light box). The "head and hand" condition was thus exactly the same as the hands-free condition of Gergely, Bekkering, and Király (2002), with the only difference being that the model demonstrated the prepotent hand touch action as well, and both actions resulted in activating the light box. The head touch action was identical to the one employed in the previous experiments, while the hand action consisted of touching the top of the lamp and lighting it up by hand. Both actions were modeled twice, in alternating order, with the first action counterbalanced across participants.

Test phase. The test phase in this condition was similar to the test phase of the communicative context condition in study 1.

Results

Briefly, in the "head and hand" condition, only 8 percent of infants (one child) imitated the head touch (see figure 5.2). We compared the results of the present study to the results of

the two conditions of study 1. The frequency of target action reenactment was lower in the "head and hand" condition than it was in the second-person "communicative demonstration context" condition of study 1 (Fisher exact $p = 0.001$ and $OR = 48.1$), and the frequency of target actions did not differ significantly between the third-person "noncommunicative observation context" condition of study 1 and the "head and hand" condition (Fisher exact $p = 0.148$).

Discussion

The results of the "head and hand" condition are in line with the assumption of inference-based learning: it was ostensively demonstrated to infants that the efficient hand action is an established, socially sanctioned, and efficient way to attain the goal. Therefore, they acquired this means action, which they judged as efficient, socially shared, and relevant. At the same time, they also saw an ostensive manifestation of an opaque alternative way to attain the same goal (the head touch action). Since it was demonstrated to infants that both means (hand touch action and head touch action) are equally relevant and socially sanctioned ways to attain the same goal, they were free to choose between them. The results show that in this case they chose to perform the more efficient means action and did not reenact the opaque (subefficient and thus more costly) alternative. Note that this finding is hard to reconcile with theories according to which "children imitate behavior that is causally opaque with higher fidelity than behavior that has a transparent physical causal mechanism" (Legare 2019, 130; also see Legare et al. 2015). Moreover, the lack of imitation of the opaque head action in the present study is hard to explain with the variety of theories of cultural transmission that consider opacity of cultural actions as automatically inducing high-fidelity behavior-copying (e.g., Tennie, Call, and Tomasello 2009).

Conclusion

In our view, to provide an adequate explanatory model of the role of emulative reenactment in human cultural learning, any viable theory must be able to account for two significant empirical properties of the way human infants acquire novel skills from observing them performed by others. First, one must account for our remarkable species-unique ability for social learning that allows even preverbal infants to learn quickly, retain over the long term, and delay functionally appropriate reenactments of novel means actions. Indeed, even in cases when a skill has been presented to them only on a single occasion and its reenactment takes place weeks or even months later—for example, as demonstrated in Meltzoff (1988, 1996)—infants are able to acquire culturally relevant skills with remarkable success.

Second, it is crucial to account for the adaptive ability of human infants—and, more widely, of "human cultural novices"—to flexibly but appropriately generalize and selectively reproduce newly acquired motor skills across a variety of functionally relevant and novel contexts. As demonstrated in our studies, the proposed relevance-based inferential account can provide solutions to these two problems since infants can encode a novel behavior after only a few demonstrations. Furthermore, our studies provide evidence that infants' functionally appropriate application of the head touch action is in response to the specific demands of the different situations. These properties of inference-based selective emulative learning,

however, represent challenges for the rigid behavior-copying model of imitative learning that accompanies the dichotomy-based approach.

The evidence presented in this chapter contradicts the basic assumptions of the dichotomy-based approach of cultural transmission by showing that when infants and young children observe demonstrations of novel goal-directed instrumental actions that contain causally opaque components and subgoals, the naive learner does not rigidly produce faithful motor replicas of the observed target action. On the contrary, the demonstration of causally opaque instrumental actions induces goal-emulative variability in reproduction.

A closer look at the concrete form of reenacted target actions uncovered how infants reproduced the action means in a remarkably flexible manner, freely generating alternative action variants. Infants did not always bend forward to contact the lamp with their forehead; instead, they either lifted the lamp up by hand or bent forward to approach the lamp with their head. Moreover, in many cases (30% in study 1 and 58% in study 2), even for infants who did not actually bring about the desired outcome, their performed actions clearly indicated their intention to achieve the observed goal. Hence, the main findings of the presented studies support the view that action understanding and goal inference precede, rather than follow, action mirroring processes (Csibra 2007). We acknowledge, however, that the studies presented in support of this general conclusion are restricted in scope and test the effect of causal opacity on cultural transmission processes only in the domain of transitive goal-directed instrumental actions—where the presence of causally opaque action components and subgoals have received less attention in theories of cultural transmission.

According to our interpretation of relevance-based goal emulation, infants encode the goal of the action contextually, and they retrieve a behavior that is effective in its attainment. In addition, it is proposed that natural pedagogy modulates what is learned in the situation (Gergely and Csibra 2006, 2009; Király, Csibra, and Gergely 2013). Ostensive communicative demonstrations can enrich the encoding of the goal by manifesting the particular means (or features thereof) as a culturally relevant manner to attain the goal, thereby indicating to the infant that the manifested action variant is worth acquiring. The presumption of relevance also induces infants' attention to the contextual factors in which the action variant is manifested, leading children to generate emulative variants to discover whether or not there is space for behavioral refinements. Finally, the presumption of relevance also leads infants to differentiate the instrumental from the conventional functions that are served by an action.

The role of ostension is to highlight a behavior, or an aspect of behavior, that is not causally accessible to the observer in that it is not analyzable with the individual's cognitive toolkit. However, ostensive demonstration induces the reproduction of behavior while allowing for variability in its reproduction. Indeed, ostensive demonstration brings into focus aspects of a novel action sequence to be encoded as subgoals that would otherwise remain opaque or unattended to by the learner. These subgoals are highlighted within specific contexts that help ground their relevance. However, being goals themselves, these subgoals can also be attained through variable means actions. This results in the fact that some subefficient means actions do not overwrite or rule out alternatives but rather often coexist with more efficient variants that are also acceptable in everyday situations (Altinok et al. 2020).

Our head touch studies underline the fact that subefficient means actions are used together with efficient alternatives. In the context of teaching, ostensive demonstration not only

induces the encoding of a relevant novel aspect of an action as a subgoal of that action, but it also allows infants to segment the action and analyze the situation of its enactment, thereby promoting learning about other specific contexts in which the novel action could be applied as a relevant means. Changes in contexts can be responsible for variability in behavior and, at the same time, contribute to the survival of both efficient and subefficient yet socially determined and conventional formats of action. Ostensive demonstrations inform the learner about how to adapt their actions in a context-dependent way.

How and why do less efficient cultural versions of means actions survive? First, as seen in our illustrative head touch studies, when children try to reach a goal, they vary the means employed in a trial-and-error manner. This process can result in more flexibility on the part of the learner, who can compare alternative versions of an action—including the versions tried out by a model—for attaining the subgoal itself (e.g., the varying way that infants touched the lamp with different parts of their head). This potential for monitoring the variants and their success during reenactment, and for relating them to different features of the context in which they tend to be produced, can also facilitate the emergence of new variants and combinations of behaviors.

This process could also result in a deeper and more detailed understanding of the different kinds of functional determinants that are involved in relating the subgoals to the final goal in particular contexts. In particular, by comparing and analyzing the alternative variants produced along the lines of efficiency and relating them to differences of the situational context in which they are more likely to appear, the learner could differentiate the relevant contributions of causal and social conventional functions in the use of alternative action variants to achieve a subgoal. Consequently, during the reproduction of the behavior in a context-dependent way, new variants could emerge. These new variants could result in new solutions and so could also fuel innovative processes (for a similar proposal, see Yu and Kushnir 2020).

If the primary mechanism of social learning were to copy blindly behaviors that appear causally opaque to the learner, there would be no room for instrumental refinement. Based on the empirical findings presented above, we suggest that the central assumption of the dichotomy-based approach—namely, that the apparent causal opacity of novel actions observed induces a rigid behavior-copying mechanism that produces faithful motor replicas and inhibits the production of variability in action alternatives—is fundamentally mistaken both when applied to the domain of noninstrumental ritual actions serving social functions and in the domain of transitive goal-directed actions associated with instrumental functions.

We propose that repeated ostensive demonstrations of causally opaque and subefficient actions aimed at attaining subgoals in certain types of social contexts function to maintain and stabilize the normative use of less efficient versions and to safeguard against their replacement in the cultural repertoire by new, successful, and more efficient alternatives. Ostensive manifestations highlight how the utilization of a subefficient action version is not accidental but instead serves a social function, which constitutes a culturally sanctioned alternative. This picture thus suggests that the ostensive demonstration of opaque action variants induces in the cultural learner a presumption of relevance and leads to relevance-based emulation of alternative variants. This source of reproductive variability in the domain of instrumental actions could foster the emergence of technological innovations by discovering more adaptive action variants or their combinations and generalizing the acquired func-

tional skill over a broad range of contexts. Even in the case of subgoals perceived as conventional, relevance-based emulation can introduce variability and could therefore allow the emergence of more efficient alternatives.

Yet, as the studies discussed by Nicola Cutting (this volume) highlight, children are poor tool innovators. While children have an outstanding ability to quickly learn the use of a tool when following the demonstration of others, it is rare for them to solve problems by designing (even simple) new tools. Cutting emphasizes how innovation is likely to be more socially mediated and how previous studies lack attention to this social aspect. What is missing for children in such cases is the ability to activate the relevant knowledge that they have already learned and reorganize it in a novel way for functional and efficient goal attainment. A recent investigation brought into focus how in hunter-gatherer societies, some socialization practices can be observed that boost innovation in children (Lew-Levy et al. 2020). These socialization practices include the support of learning through autonomous exploration (see Boyette, this volume), the teaching of children by adults and peer-play, and sensitizing children for seeking novel forms of goal attainment (innovation seeking). Genuine innovation, consequently, builds on socially mediated, accumulated knowledge as well as the capacity to introduce variations. That is why innovative tool design seems to be a late-developing competence. In this sense, the flexibility of social learning is a prerequisite to the emergence of innovative capabilities.

To summarize, the view that we propose here is the following. On the one hand, relevance-based goal emulation is a social learning mechanism that promotes the discovery of variants in behavior through the understanding of goals and subgoals organized into hierarchies within an action. On the other hand, this mechanism contributes to the establishment of a robust and rich knowledge base by facilitating the production of alternative variants in different contexts. In our view, these two aspects actually represent different forms of learning carried out by the same mechanism—that is, relevance-based emulation—that are jointly necessary for achieving innovation.

Notes

1. When one observes a behavior, one employs mental causal models when interpreting it. A behavior is causally transparent when the available contextual information allows for interpretation of its causal structure using such model. For instance, when an agent makes a detour to reach a goal object, this behavior is causally transparent if one sees an obstacle to go around. In contrast, a behavior is causally opaque when the contextual information is insufficient to interpret the behavior's causal structure—for example, when an agent makes a detour to reach a goal object, this behavior is difficult to understand if there is no obstacle to go around (see Gergely, Bekkering, and Király 2002).

2. Note, however, that by identifying the causally unnecessary transitive actions as being "silly," one suggests that the children recognized them as being causally transparent, as indicated by their correct judgment that they were causally irrelevant for retrieving the goal object. Despite this understanding, these actions were faithfully copied and reproduced just as were the equally unnecessary but causally opaque intransitive "magic" gestures that were also part of the demonstrated series of actions.

3. The term "innovation" is defined by proponents of the dichotomy-based view of cultural transmission as "constructing new tools, or using old tools in new ways, to solve new problems" (see Legare and Nielsen 2015, 689).

4. Some forms of everyday behaviors appear causally opaque even for skilled adults, who habitually eat food in the normative manner of their own cultural tradition or do so selectively as a function of the particular social context, such as at a formal dinner with the president of a university.

References

Altınok, N., M. Hernik, I. Király, I., and G. Gergely. 2020. "Acquiring Sub-efficient and Efficient Variants of Novel Means by Integrating Information from Multiple Social Models in Preschoolers." *Journal of Experimental Child Psychology* 195:104847.

Altınok, N., I. Király, and G. Gergely. 2022. "The Propensity to Learn Shared Cultural Knowledge from Social Group Members: Selective Imitation in 18-Month-Olds." *Journal of Cognition and Development* 23 (2): 273–288.

Berl, R. E., and B. S. Hewlett. 2015. "Cultural Variation in the Use of Overimitation by the Aka and Ngandu of the Congo Basin." *PLOS One* 10 (3): e0120180.

Boyd, R., and P. J. Richerson. 1996. "Why Culture Is Common, but Cultural Evolution Is Rare." In *Evolution of Social Behaviour Patterns in Primates and Man*, edited by W. G. Runciman, J. M. Smith, and R. I. M. Dunbar, 77–93. Oxford: Oxford University Press.

Buchsbaum D., A. Gopnik, T. L. Griffiths, and P. Shafto. 2011. "Children's Imitation of Causal Action Sequences Is Influenced by Statistical and Pedagogical Evidence." *Cognition* 120 (3): 331–340.

Carpenter, M., J. Call, and M. Tomasello. 2005. "Twelve- and 18-Month-Olds Copy Actions in Terms of Goals." *Developmental Science* 8 (1): 13–20.

Chen, M., I. Király., and G. Gergely. 2012. "*Opacity and Relevance in Social Cultural Learning: Relevance-Guided Emulation in 14-Month-Olds*." Paper presented at Budapest CEU Conference on Cognitive Development, Budapest, January 12–14.

Clegg, J. H., and C. H. Legare. 2016. "Instrumental and Conventional Interpretations of Behavior Are Associated with Distinct Outcomes in Early Childhood." *Child Development* 87 (2): 527–542.

Csibra, G. 2007. "Action Mirroring and Action Understanding: An Alternative Account." In *Sensorimotor Foundations of Higher Cognition: Attention and Performance XXII*, edited by P. Haggard, Y. Rosetti, and M. Kawato, 435–458. Oxford: Oxford University Press.

Csibra, G., and G. Gergely. 2006. "Social Learning and Social Cognition: The Case for Pedagogy." In *Processes of Change in Brain and Cognitive Development: Attention and Performance XXI*, edited by Y. Munakata and M. H. Johnson, 249–274. Oxford: Oxford University Press.

Csibra, G., and G. Gergely. 2009. "Natural Pedagogy." *Trends in Cognitive Sciences* 13 (4): 148–153.

Csibra, G., and G. Gergely. 2011. "Natural Pedagogy as Evolutionary Adaptation." *Philosophical Transactions of the Royal Society B: Biological Sciences* 366 (1567): 1149–1157.

Gergely, G. 2007. "The Social Construction of the Subjective Self: The Role of Affect-Mirroring, Markedness, and Ostensive Communication in Self-Development." In *Developments in Psychoanalysis. Developmental Science and Psychoanalysis: Integration and Innovation*, edited by L. Mayes, P. Fonagy, and M. Target, 45–88. London: Karnac Books.

Gergely, G., H. Bekkering, and I. Király. 2002. "Rational Imitation in Preverbal Infants." *Nature* 415 (6873): 755–755.

Gergely, G., and G. Csibra. 2003. "Teleological Reasoning in Infancy: The Naïve Theory of Rational Action." *Trends in Cognitive Sciences* 7:287–292.

Gergely, G., and G. Csibra 2006. "Sylvia's Recipe: The Role of Imitation and Pedagogy in the Transmission of Human Culture." In *Roots of Human Sociality: Culture, Cognition, and Human Interaction*, edited by N. J. Enfield and S. C. Levinson, 229–255. Oxford: Berg Publishers.

Gergely, G., and P. Jacob. 2012. "Reasoning about Instrumental and Communicative Agency in Human Infancy." In *Rational Constructivism in Cognitive Development*, edited by F. Xu and T. Kushnir, 59–94. New York: Academic Press.

Gergely, G., Z. Nádasdy, G. Csibra, and S. Bíró. 1995. "Taking the Intentional Stance at 12 Months of Age." *Cognition* 56 (2): 165–193.

Heintz, C., and T. Scott-Phillips. 2023. "Expression Unleashed: The Evolutionary and Cognitive Foundations of Human Communication." *Behavioral and Brain Sciences* 46 (January): 1–46.

Henrich, J., S. J. Heine, and A. Norenzayan. 2010. "The WEIRDest People in the World?" *Behavioral and Brain Sciences* 33 (2–3): 61–83.

Hewlett, B. S., and C. J. Roulette. 2016. "Teaching in Hunter-Gatherer Infancy." *Royal Society Open Science* 3:150403.

Hoehl, S., S. Keupp, H. Schleihauf, N. McGuigan, D. Buttelmann, and A. Whiten. 2019. "'Over-Imitation': A Review and Appraisal of a Decade of Research." *Developmental Review* 51:90–108.

Kapitány, R., and M. Nielsen. 2019. "Ritualized Objects: How We Perceive and Respond to Causally Opaque and Goal Demoted Action." *Journal of Cognition and Culture* 19:170–194.

Király, I., G. Csibra, and G. Gergely. 2013. "Beyond Rational Imitation: Learning Arbitrary Means Actions from Communicative Demonstrations." *Journal of Experimental Child Psychology* 116 (2): 471–486.

Lancy, D. F. 1996. *Playing on the Mother-Ground: Cultural Routines for Children's Development*. New York: Guilford Press.

Lancy, D. F., S. Gaskins, and J. Bock, eds. 2009. *The Anthropology of Learning in Childhood*. Lanham, MD: Altamira Press.

Legare, C. H. 2019. "The Development of Cumulative Cultural Learning." *Annual Review of Developmental Psychology* 1 (1): 119–147.

Legare, C. H., and M. Nielsen. 2015. "Imitation and Innovation: The Dual Engines of Cultural Learning." *Trends in Cognitive Sciences* 19 (11): 688–699.

Legare, C. H., and M. Nielsen. 2020. "Ritual Explained: Interdisciplinary Answers to Tinbergen's Four Questions." *Philosophical Transactions of the Royal Society B: Biological Sciences* 375 (1805): 1–5.

Legare, C. H., N. J. Wen, P. A. Herrmann, and H. Whitehouse. 2015. "Imitative Flexibility and the Development of Cultural Learning." *Cognition* 142:351–361.

Lew-Levy, S., A. Milks, N. Lavi, S. M. Pope, and D. E. Friesem. 2020. "Where Innovations Flourish: An Ethnographic and Archaeological Overview of Hunter-Gatherer Learning Contexts." *Evolutionary Human Sciences* 2 (2020): e31.

Lyons, D. E., A. G. Young, and F. C. Keil. 2007. "The Hidden Structure of Overimitation." *Proceedings of the National Academy of Sciences* 104 (50): 19751–19756.

McGuigan, N., A. Whiten, E. Flynn, and V. Horner. 2007. "Imitation of Causally Opaque versus Causally Transparent Tool Use by 3- and 5-Year-Old Children." *Cognitive Development* 22 (3): 353–364.

Meltzoff, A. N. 1988. "Infant Imitation after a One-Week Delay: Long-Term Memory for Novel Acts and Multiple Stimuli." *Developmental Psychology* 24 (4): 470–476.

Meltzoff, A. N. 1996. "The Human Infant as Imitative Generalist: A 20-Year Progress Report on Infant Imitation with Implications for Comparative Psychology." In *Social Learning in Animals: The Roots of Culture*, edited by C. M. Heyes and B. G. Galef, 347–370. New York: Academic Press.

Nielsen, M., and K. Tomaselli. 2010. "Overimitation in Kalahari Bushman Children and the Origins of Human Cultural Cognition." *Psychological Science* 21 (5): 729–736.

Over, H., and M. Carpenter. 2012. "Putting the Social into Social Learning: Explaining Both Selectivity and Fidelity in Children's Copying Behavior." *Journal of Comparative Psychology* 126 (2): 182–192.

Over, H., and M. Carpenter. 2013. "The Social Side of Imitation." *Child Development Perspectives* 7:6–11.

Paradise, R., and B. Rogoff. 2009. "Side by Side: Learning by Observing and Pitching In." *Ethos* 37:102–138.

Paulus, M., S. Hunnius, M. Vissers, and H. Bekkering. 2011. "Imitation in Infancy: Rational or Motor Resonance?" *Child Development* 82:1047–1057.

Pinkham, A. M., and V. K. Jaswal. 2011. "Watch and Learn? Infants Privilege Efficiency over Pedagogy during Imitative Learning." *Infancy* 16:535–544.

Rendell, L., L. Fogarty, W. J. Hoppitt, T. J. Morgan, M. M. Webster, and K. N. Laland. 2011. "Cognitive Culture: Theoretical and Empirical Insights into Social Learning Strategies." *Trends in Cognitive Sciences* 15 (2): 68–76.

Sperber, D., and D. Wilson. 2002. "Pragmatics, Modularity and Mind-Reading." *Mind & Language* 17 (1–2): 3–23.

Tennie, C., J. Call, and M. Tomasello. 2009. "Ratcheting Up the Ratchet: On the Evolution of Cumulative Culture." *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1528): 2405–2415.

Tennie, C., and C. P. van Schaik. 2020. "Spontaneous (Minimal) Ritual in Non-human Great Apes?" *Philosophical Transaction of the Royal Society B: Biological Sciences* 375:20190423.

Toelch, U., D. R. Bach, and R. J. Dolan. 2014. "The Neural Underpinnings of an Optimal Exploitation of Social Information under Uncertainty." *Social Cognitive and Affective Neuroscience* 9 (11): 1746–1753.

Tomasello, M. 1996. "Do Apes Ape?" In *Social Learning in Animals*, edited by C. M. Heyes and B. G. Galef, 319–346. New York: Academic Press.

Tomasello, M. 1998. "Emulation Learning and Cultural Learning." *Behavioral and Brain Sciences* 21 (5): 703–704.

Tomasello, M. 1999. *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

Watson-Jones, R. E., and C. H. Legare. 2016. "The Social Functions of Group Rituals." *Current Directions in Psychological Science* 25 (1): 42–46.

Wen, N. J., P. A. Herrmann, and C. H. Legare. 2016. "Ritual Increases Children's Affiliation with In-Group Members." *Evolution and Human Behavior* 37 (1): 54–60.

Yu, Y., and T. Kushnir. 2020. "The Ontogeny of Cumulative Culture: Individual Toddlers Vary in Faithful Imitation and Goal Emulation." *Developmental Science* 23 (1): e12862.

This is a section of doi:10.7551/mitpress/15181.001.0001

The Evolution of Techniques

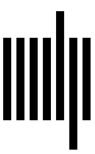
Rigidity and Flexibility in Use, Transmission, and Innovation

Edited by: Mathieu Charbonneau

Citation:

The Evolution of Techniques: Rigidity and Flexibility in Use, Transmission, and Innovation
Edited by: Mathieu Charbonneau
DOI: 10.7551/mitpress/15181.001.0001
ISBN (electronic): 9780262378390
Publisher: The MIT Press
Published: 2024

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open



The MIT Press

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Times New Roman by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data is available.

ISBN: 978-0-262-54780-2