The social construction of the cultural mind

Imitative learning as a mechanism of human pedagogy

György Gergely and Gergely Csibra
Institute for Psychological Research, Hungarian Academy of Sciences, Budapest / Centre for Brain and Cognitive Development, Birkbeck College, London

How does cultural knowledge shape the development of human minds and, conversely, what kind of species-specific social-cognitive mechanisms have evolved to support the intergenerational reproduction of cultural knowledge? We critically examine current theories proposing a human-specific drive to *identify with* and *imitate* conspecifics as the evolutionary mechanism underlying cultural learning. We summarize new data demonstrating the *selective interpretive nature of imitative learning* in 14-month-olds and argue that the predictive scope of existing imitative learning models is either too broad or too narrow to account for these findings. We outline our alternative theory of a human-specific adaptation for 'pedagogy', a communicative system of mutual design specialized for the fast and efficient transfer of new and relevant cultural knowledge from knowledgeable to ignorant conspecifics. We show the central role that innately specified ostensive-communicative triggering cues and learner-directed manner of knowledge manifestations play in constraining and guiding selective imitation of relevant cultural knowledge that is both new and cognitively opaque to the naive learner.

Keywords: imitative learning, social learning, social cognitive development, cultural learning, early teleological reasoning, pedagogical stance

Introduction: Imitative learning as a human-specific adaptation for cultural transmission

Minds construct culture and culture constructs minds. The ontogenetic development of the human mind is deeply influenced both by the characteristics of

ISSN 1572–0373 / E-ISSN 1572–0381

the multitude of cultural products it encounters, and by the relevant behaviours of their knowledgeable users that it observes. But the reverse, we shall argue, is also true: forms of human culture would not be able to spread and survive cross-generationally had the mind of the human infant not been equipped with adapted cognitive resources specialized for the reception and transmission of relevant cultural knowledge. Therefore, one of the central issues raised by human culture concerns the nature of the social-cognitive mechanisms that mediate the reproduction, spread, and intergenerational transmission of cultural forms among members of the community.

The dominant candidate for such a mechanism has long been the special human capacity and inclination to *imitate* the actions of conspecifics. *Imitative learning* has been proposed as a human-specific adaptation for cultural learning (e. g., Meltzoff, 1996; Tomasello 1999; Tomasello et al., 1993; 2005) for several reasons. First, humans — more than most other species — are prolific and flexible imitators, who seem specially adapted to imitate a wide range of behaviors, often without direct reinforcement (Meltzoff, 1996). Second, while 'cultural' behavioral traditions (such as group-specific termite fishing or nut cracking techniques) also exist in non-human primates (Goodall, 1986; Whiten et al., 1999), it has been argued that such cultural skills are socially transmitted through observational learning mechanisms that do *not* involve imitation (such as stimulus enhancement, response facilitation, or trial-and-error emulation) (Heyes & Galef, 1996; Thorpe, 1963; Tomasello, 1996; Tomasello & Call, 1997).

In this paper we shall critically reexamine the dominant role attributed to imitative learning in the intergenerational transmission of human cultural knowledge. First, we shall consider two influential recent proposals (one by Andy Meltzoff (1996), the other by Mike Tomasello and colleagues (1993, 2005)) claiming that an identification-based drive to imitate the actions of conspecifics forms the central species-specific adaptation for cultural learning in humans. We shall evaluate these theories in the light of new evidence (Gergely, Bekkering, & Király, 2002; Király, Csibra, & Gergely, 2004) demonstrating the selective interpretive nature of imitative learning in human infants. It will be argued that the predictive scope of the two theories in question is either too broad (Meltzoff) or too narrow (Tomasello) to account for the relevance-based selectivity that characterizes young infants' imitative learning of novel means. We shall contrast these models with our own alternative proposal for a human-specific adaptation for 'pedagogy', a complex communicative system of mutual design specialized for the fast and efficient transmission of cultural knowledge

(Csibra and Gergely, 2005; Gergely and Csibra, 2005a). We shall argue that imitation is not an adaptation for human cultural learning in its own right, but only a basic low-level capacity (available to many non-human species (Heyes, 1993)) that in humans have become recruited as a subcomponent of the system of pedagogical knowledge transfer. In closing, we shall show how the selective, relevance-guided nature of early imitative learning can be best accounted for as a result of the constraining effects of the built-in assumptions of the 'pedagogical stance' about the ways in which relevant cultural information is ostensively communicated by knowledgeable others for the sake of naive conspecifics. These assumptions function in humans to guide and constrain imitative learning by identifying the culturally relevant contents for the learner to be retained and fast learned.

Imitative learning as a human-specific drive to "act like" other humans

Based on their demonstrations of neonatal imitation, Meltzoff and Moore (1977, 1989, 1997) argued that (a) human infants have a prewired mechanism to map observed behavior of others onto the corresponding motor scheme of the self, (b) this mechanism allows infants to recognize others as conspecifics, as being "just-like-them", (c) infants have an innate predisposition to "identify" with others perceived as "just-like-them", and (d) they "have an inbuilt drive to "act like" their conspecifics" (Meltzoff, 1996, p. 363).

In a seminal study, Meltzoff (1988) has shown that this innate propensity to imitate humans also leads infants very early on to *imitatively learn novel means actions* from observing others. Fourteen-month-olds watched as a human model illuminated a magic light-box by leaning forward from waist and touching its top panel with her forehead. A week later, 67% of the infants reenacted the novel 'head-action', while none performed it in a base-line control group that had not seen the action demonstrated. This illustrates how, in Meltzoff's theory, the infant's innate drive for identification, and the consequent tendency to imitate other humans, also provide the basic mechanism for cultural learning.

Imitation and mindreading: "Insightful" imitation as a precondition for cultural learning

Imitative learning also plays a crucial role in Tomasello's (1999; Tomasello et al., 1993, 2005) theory of cultural learning. Similarly to Meltzoff, Tomasello (1999) suggests that "Imitative learning… relies fundamentally on infants' tendency to *identify* with adults" (p. 82). Tomasello et al. (2005) argue, however, that simple "bodily identification" as evidenced by neonatal imitation is not sufficient to support the kind of imitative learning that makes human cultural learning possible. For Tomasello, 'true' imitative learning *qua* cultural learning also necessitates the understanding of the *intentions* behind the other's action. "True imitative learning…involves the infant's reproducing the adult's actual behavioral strategies in their appropriate functional contexts, which implies an understanding of the intentional state underlying the behavior" (Tomasello et al., 1993, p. 497). Apart from this strong cognitive requirement for imitative learning, Tomasello et al. (2005) also postulate a uniquely human motivational precondition in the form of a primary human-specific "motivation to share psychological states with others" (p. 1), which leads to "more deeply psychological levels of identification" (p. 26).[1] It is at this level of *identification* with the mental states of conspecifics that infants can *simulate* the other's intentional actions, attributing the simulated intention automatically to the other (Tomasello, 1999, pp. 73–76).

Tomasello (1999) argues that apes lack the capacity for identification, therefore, they don't simulate and attribute intentions to others either. In his view, this is reflected also in the fact that apes seem not to learn from observing others through imitation, but only through trial-and-error 'emulation' (Tomasello, 1996): they try to reproduce the observed outcome in their own way without attending to or directly re-enacting the particular means action observed. Eventually, through this slow process of (re)discovery, apes manage to acquire the same skill that they observed to produce the desired outcome (or some variant of it).

If one is simply 'blindly copying' an action without understanding the intention behind it, one cannot speak about 'true' imitative learning either, only about "mimicry". Therefore, Tomasello (1996; Tomasello et al., 1993) provided criteria for 'true' imitative learning to help differentiate it from 'pseudo-imitative' re-enactments of others' behaviors produced by 'emulation' or 'mimicry', social learning processes that are also available to many non-human animals.

1. *"The novelty-of-response criterion"*: To avoid confusion with response facilitation, imitative learning should involve "the learning of a new response" (Byrne and Tomasello, 1995) that is not part of the organism's motor repertoire.
2. *"The cognitive transparency criterion"*: Apart from providing a safeguard against confusing imitative learning with 'blind mimicry', this requirement also functions as a *'selection filter'* determining "which aspects of the behavior are relevant for reproduction" (Tomasello, 1996, p. 323). Imitative learning entails "an understanding of both the behavior's goal and its strategy for achieving that goal" (p. 324). To imitatively learn a novel behavioral strategy, the infant must understand "…how the behavior is designed to bring about the goal. This then determines precisely what of the other's behavior it seeks to reproduce" (p. 324). In other words, according to Tomasello we can only speak of 'true' imitative learning, when the underlying causal intentional structure of the other's imitated action is 'cognitively fully transparent' to the learner.

Let us now revisit Meltzoff's (1988) 'magic-box' experiment with Tomasello's criteria in mind. First, the fact that most infants imitated the novel 'head-action', "an unusual and awkward behavior…even though it would have been easier and more natural for them simply to push the panel with their hand" (Tomasello, 1999, p. 82), rules out an emulation account and satisfies Tomasello's "novelty-of-response criterion". Want and Harris (2002), having similarly ruled out emulation, argued, however, that "the children seem to have mimicked or 'blindly' imitated the demonstration, copying exactly the actions…demonstrated" (p. 8).

But 'blind' mimicry implies that the organism copies the behavior "without any regard for its goal-directed nature" (Tomasello, 1999, p. 82). To test this, Carpenter, Nagell, and Tomasello (1998) replicated the Meltzoff (1988) task so that the 'head-action' and its outcome were spatially separated (the light-source was above the box). They found that "the majority of infants both reproduced the unusual action and looked to the interesting result in anticipation — demonstrating that they were not just mimicking" (Tomasello, 1999, p. 82).

Note, however, that this is a rather "weak" test of Tomasello's (1996) full "cognitive transparency criterion" according to which 'true' imitative learning should entail "an understanding of both the behavior's goal *and its strategy for achieving that goal*" (p. 324, emphasis added). Even more strongly, 'true' imitative learning takes place only if the infant understands "…*how the behavior is designed to bring about the goal*. This then *determines precisely* what of the oth-

er's behavior it seeks to reproduce" (p. 324, emphasis added). However, it seems entirely doubtful that the 14-month-olds could have reconstructed — through simulation — the underlying intention and rational design behind the model's choice to perform the 'head-action' rather than the apparently more sensible, efficient, and readily available 'hand-action'. Strictly speaking, "the cognitive transparency criterion" should have predicted that infants will *not* imitate the bizarre 'head-action', as it must have remained cognitively opaque to them in terms of the actor's underlying reasons for performing the — apparently non-optimal — means action.

Teleological emulation *versus* rational imitation: The selective interpretive nature of imitative learning in human infants

Meltzoff's (1988) finding that 14-month-olds readily imitate the unusual 'head-action', seemed also unexpected from the point of view of our own theory of the one-year-old's "teleological stance" or "naïve theory of rational action" (Csibra & Gergely, 1998; Gergely & Csibra, 2003). In a series of violation-of-expectation looking time studies (Gergely, Nádasdy, Csibra, & Bíró, 1995; Csibra, Bíró, Koós, & Gergely, 2003), we have shown that by 12 months infants exhibit a sophisticated ability to attribute goals to observed actions and to evaluate the relative efficiency of the means act in relation to the goal and the physical constraints of the actor's situation. If they know the actor's goal and see a change in situational constraints, young infants can infer what the most efficient new means would be to the goal in the new situation and expect that the actor 'ought to' perform that particular action to achieve the goal[2] (Gergely & Csibra, 2003). On that ground, then, one would have expected that in the Meltzoff (1988) task infants, as rational agents, should have performed the most efficient goal-directed action available to them (using their hand to contact the light-box), instead of imitating the awkward and less efficient 'head-action'.

To clarify this situation, Gergely, Bekkering, and Király (2002) performed a modified version of the Meltzoff (1988) task. They hypothesized that "if infants noticed that the demonstrator declined to use her hands despite the fact that they were free, they may have inferred that the head action must offer some advantage in turning on the light. They therefore used the same action themselves in the same situation" (p. 755). To test this idea, Gergely et al. ran two groups of 14-month-olds varying the situational constraints of the model. In the 'Hands-occupied' condition the model's hands were visibly occupied: she pretended

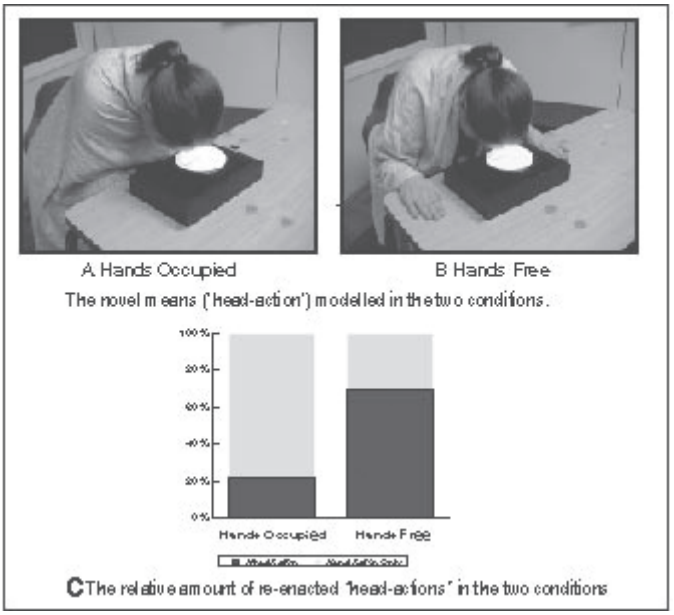

Figure 1.

to be chilly and wrapped a blanket around her shoulders holding it with both hands while performing the 'head-action'. In the 'Hands-free' condition, however, after wrapping the blanket around her shoulders, the model placed her hands visibly free onto the table before demonstrating the 'head-action'.

As Figure 1 shows, when the model's hands were occupied, 14-month-olds were much less likely to imitate the 'head-action' (21%). Instead, they illuminated the box by *touching it with their hand* performing the most sensible, simpler, easy-to-perform, and equally effective emulative response available to them, but not to the model (teleological emulation). In contrast, when the model's hands were free, but she still used her head to illuminate the box, 69% of 14-month-olds imitated her 'head-action' ($p < .02$) (replicating Meltzoff, 1988).

A further unexpected finding was that in *both* conditions *all* infants performed the emulative 'hand-action'. Moreover, all subjects in the 'Hands-free' condition who imitated the 'head-action', did so only *after* they had first performed the 'hand-action' that — in all cases — succeeded in illuminating the box. In other words, even after they have experienced that the effect can be brought about by the simpler 'hand-action' as well, most infants in the 'Hands-free' condition remained motivated to imitate the model's demonstrated — though apparently less efficient — 'head-action'.

Let us draw some preliminary conclusions for the two theories of imitative learning described above. First, our findings suggest that imitative learning of novel means is not triggered by identification (as on that basis one could not have predicted a significant difference in imitation between the two context-conditions). Second, our results indicate that imitative learning is not due to automatic behavioral 'copying' of the modeled action. Rather, it is guided and constrained by a top-down *selective interpretive process* involving the *evaluation of the relative efficiency of the means action* as a function of the actor's situational constraints ('Hands-free' *vs.* 'Hands-occupied').

Note, first, that Meltzoff's (1996) theory of the human infant as an "imitative generalist" driven by an innate drive to identify with and act like other humans, contains no mechanism that could account for the selectivity of imitation as a function of the actor's situational constraints. As it stands, Meltzoff's theory predicts automatic copying of *any* observed human action and so it has an *overly broad predictive scope* contradicted by the selective nature of imitative learning demonstrated.

In contrast, Tomasello's theory does include a 'selective filter' in the form of his full 'cognitive transparency criterion'. His model, however, has an *overly narrow predictive scope* as it generates wrong predictions concerning *what* will be imitated. Tomasello's theory predicts that infants will imitate only those behaviors whose underlying intentions and rational design they can fully understand through simulation. Therefore, as it stands, his theory cannot account for the imitative learning of truly *novel* behavioural means that are unpredictable on the grounds of physical-causal efficiency considerations and that, therefore, remain cognitively 'opaque' to the infant.

Cultural learning and human pedagogy

We shall now turn to our own interpretation of the nature of imitative learning and its role in the transmission of human cultural knowledge. We propose that the basic capacity to imitatively 'copy' observed behaviors of conspecifics (present in numerous non-human species as well) has evolved to serve a uniquely human function as a mechanism recruited, directed, and constrained by *pedagogy,* a specialized human-specific cognitive system dedicated to cultural learning. In our view, pedagogy was selected as a primary species-specific cognitive adaptation of mutual design to ensure fast and efficient transfer of relevant cultural knowledge through ostensive communicative 'teaching'

manifestations of relevant information by knowledgeable humans for the sake of ignorant learners (Csibra & Gergely, 2005; Gergely & Csibra, 2005a).

What may have been the evolutionary origins of human pedagogy? Elsewhere (Csibra & Gergely, 2005; Gergely & Csibra, 2005a) we speculate that during hominid evolution the original 'simple' goal-driven teleological reasoning capacity of our ancestors about objects as transient tools in the visible presence of goals (answering the question: 'What object could I use to achieve this goal?') was superseded by a more stable functionalist conceptualization in terms of affordance properties (giving rise to "inverse teleological reasoning" answering the question: 'What purpose could I use this object for?'). This eventually led to the practice of tool manufacturing in the absence of directly visible goals as well as to the appearance of mediated tool use (i.e., using tools to make other tools: "recursive teleology"). Such advanced practices posed a learnability problem for the naive juvenile observer for whom they remained cognitively 'opaque' as — lacking perceptual information about the goal — they could not identify which aspects of the observed actions were relevant (and should, therefore, be acquired) and which were incidental. Unguided forms of existing social observational learning mechanisms (including statistical, trial-and-error, and emulation learning) were ill-suited, error-prone and too slow to solve this learnability problem and could not ensure sufficiently high-fidelity successful transgenerational transmission of such cognitively 'opaque' cultural forms and skills. In statistically-based learning mechanisms the local adaptivity of the acquired behavior is ensured by reinforcement, while its evolutionary relevance is ensured by the pattern of environmental invariance it exhibits that is gradually extracted from observed repetitions of contingencies. This makes associative learning a necessarily slow and gradual process restricted to the domain of perceivable repetitive contingencies coupled with reinforcement.

Therefore, the increasing cognitive 'opacity' of complex artifacts and their manufacturing procedures may have provided selective pressure for the evolution of a qualitatively new type of social learning mechanism in the form of pedagogy. In cultural learning one obvious way to overcome the limitations of statistically-based learning mechanisms is to acquire the relevant knowledge directly from another conspecific who already possesses it. As new behaviors, especially cultural activities, are often not transparent as to either their knowledge-base or their function, an active communicative role of the more knowledgeable conspecific may greatly assist the efficient and fast transmission of such culturally relevant information. We propose (Csibra & Gergely, 2005; Gergely & Csibra, 2005a) that Mother Nature's 'trick' to make fast and efficient

learning of complex — and, for the learner, cognitively 'opaque' — cultural knowledge possible was to have humans evolve specialized cognitive resources that form a dedicated interpersonal system of mutual design in which one is predisposed to 'teach' and to 'learn' new and relevant cultural information to (and from) conspecifics. We hypothesize that humans possessing cultural knowledge are naturally inclined not only to *use*, but also to *ostensively manifest*[3] their knowledge to (and for the benefit of) naive conspecifics, while the latter are naturally motivated to acquire such knowledge by actively seeking out, attending to, and being specially receptive to the ostensive communicative manifestations of others.

In the design specifications of pedagogical knowledge transfer it is the very fact that a knowledgeable conspecific (a 'teacher') *ostensively communicates* her cultural knowledge by *manifesting* it for the novice (the 'learner') is what ensures the (cultural) relevance of the knowledge transmitted. Since the learner is predisposed to interpret the teacher's ostensive-communicative cues that accompany his knowledge manifestation (such as eye-contact, eye-brow flashing, turn-taking contingency, see Csibra & Gergely, 2005, for a review) as evidence that the manifestation will convey *new* and *relevant* cultural information for him, this allows for fast learning of the communicated content without any further need to test its relevance independently. Furthermore, the built-in presumption of relevance of pedagogically communicated knowledge manifestations also opens the door for the acquisition of knowledge contents that are not only *arbitrary, conventional*, and causally/functionally *non-transparent*, but that sometimes don't seem to (or actually do not) have any obvious adaptive value at all (these being uniquely characteristic species-specific features of many human cultural forms).

We further propose (Csibra and Gergely, 2005) that the human-specific pedagogical inclination to transmit relevant and new cultural information to conspecifics is complemented by a *special kind of receptivity* to benefit from such teaching. Human infants are equipped with specialized cognitive resources that enable them to learn from infant-directed teaching: they 1. show early sensitivity to *communicative and ostensive cues* indicating teaching contexts (such as eye-contact, contingent reactivity, motherese, and hearing one's own name), 2. tend to interpret certain directional actions (e.g., gaze-shift or pointing) occurring in these communicative contexts as *referential cues* to identify the *referents* about which new information will be provided, 3. expect the "teacher" to ostensively manifest *relevant* and *new* information about the referent, and 4. are ready to *fast-map* such information to the referent (see Csibra & Gergely, 2005,

for reviewing supporting evidence). Finally, we hypothesize that the infant's 'pedagogical stance' contains the implicit assumption that the information revealed about the referents in such ostensive-communicative teaching contexts consists of *publicly shared* and *universal cultural knowledge* that is *generalizable* and *shareable* with other members of the cultural community.[4]

Imitative learning in the service of human pedagogy: The role of ostensive-communicative cues

It is noteworthy that studies investigating early imitative learning typically present the target behaviors in a rich ostensive communicative-referential context. For example, when a model demonstrates a novel means act (as in Meltzoff, 1988), she typically first establishes *eye-contact* with the infant often also *addressing him by his name* (ostensive cues), then *shifts her eye-gaze or point* to the referent object (referential cues). This is followed by some *communicative-referential speech act* (e.g., "Look, I'll show you something!") before the target action is demonstrated. In fact, this is highly natural and representative of the manner in which human adults manifest to a child new and relevant cultural knowledge for her to acquire.

We hypothesize that in human infants imitative learning is triggered by such pedagogical cues accompanying others' manifestations of cultural information. Furthermore, we argue that the interpretive selectivity guiding what aspect of the modelled behavior will be imitatively learned is directed and constrained by the implicit assumptions of the infant's 'pedagogical stance' that the other's ostensive cues activate. When taking the 'pedagogical stance', infants interpret the other's ostensive communicative gestures as indicating that he is about to manifest 'for' them some significant aspect of cultural knowledge that will be *new* and *relevant* and that, therefore, should be fast-learned.[5]

Let us illustrate how pedagogy works by interpreting the selective imitation finding of the Gergely et al. (2002) study in terms of the inferences invoked by the pedagogical cuing context. First, we assume that 14-month-olds interpret the ostensive-communicative cues of the model as indicating that the other is about to *manifest culturally relevant and new information* for him. Second, the pedagogical context induces in the infant a special attentional and interpretive attitude to apply his knowledge-base and available interpretative capacities (his explanatory schemes or conceptual 'modes of construals') (see Keil, 1995, 2003; Kelemen, 1999a, b; Gergely and Csibra, 2003) to infer what aspect of the

manifested behavior conveys *new* and *relevant* information. Third, the pedagogical context triggers a special receptive learning mode to *fast learn* what the infant has inferred to be new and relevant information in the manifested action.

Take the 'Hands occupied' condition. Clearly, the novel outcome including the manifested affordance property of the object (illuminability-upon-contact) is *new* information previously unknown to the infant, so it is going to be retained in memory and reproduced through action. But what about the particular behavioral means ('head-action') performed? Taking the teleological stance towards actions (Gergely and Csibra, 2003) infants can infer that given the physical constraints of the actor (hands occupied), touching the box by her forehead does, in fact, qualify as a sensible, justifiable, and efficient means to the goal. So, since the physical-causal efficiency of the 'head-action' is cognitively 'transparent' (i.e., justifiable, expectable or even predictable) for the infant who sees that the actor's hands are occupied, the fact that she used her head (and not her hands) to touch the box does *not* qualify as part of the *new* information that is being conveyed. Therefore, it is predicted that the infant will *not* imitate the 'head-action' in the 'Hands-occupied' context-condition, but will reproduce the novel information (will illuminate the box) by the most efficient means available to him given his *own* situational constraints: i.e., he will use his (free) hands to illuminate the box.

In the 'Hands-free' condition the situation is different, however. Of course, the goal-state involving the newly experienced affordance of the box is *new* information here, too, so it will be retained and reproduced. In contrast, when setting up a teleological interpretation as to what particular action would constitute under these situational constraints the most rational/efficient means to the goal, given the fact that the actor's hands were free, the infant must have identified the available 'hand-action' as the most efficient (and, therefore, expectable) means that the model 'ought to' perform. Unexpectedly, however, the demonstrator chose not to use her free hands, but performed the unusual 'head-action' instead. We hypothesize that this *perceived mismatch* between the predictable and the actually performed means drew the infants' attention to the *model's contrastive choice* to perform the unexpected 'head-action' as carrying special communicative significance. This contrastive choice then "marked" the 'head-action' as also forming part of the *new* and *relevant* information that the ostensive-communicative manifestation conveyed. As a result, both the new goal *and the new means* were retained and imitated!

Notice the contrast between this analysis and Tomasello's 'cognitive transparency criterion' that predicted that infants would imitatively learn a new

action only if its underlying intentions are cognitively fully 'transparent' and interpretable for the infant. On the contrary, our proposed pedagogical interpretation of the ostensive-communicative cues accompanying the action manifestation suggests that imitative learning of a new behavior will occur *precisely* when the choice or particular manner of the action is *unpredictable — i.e., cognitively 'opaque' — to the infant* and, as such, it qualifies as part of the 'new and relevant' information manifested by the ostensive 'teaching' act. (Note that in this model no cognitive 'insight' into *what* makes the manifested skill culturally relevant — apart from it being ostensively manifested — is presumed to be necessary for imitative learning to take place.)

But would then *any behavior that is unpredictable* lead to imitation if presented in a pedagogical context? Well, as it turns out the answer is 'no' and Tomasello is probably right in his intuition that some level of 'cognitive interpretability' is necessary for imitative learning. We have recently run a 'no-effect' control ('Hands-free' condition) in which the same ostensive cues introduced an identical 'head-bending action' without, however, the demonstrator's head actually contacting the box (it stopped 10 cm above it). So the behavior resulted in no observable external effect. In stark contrast to the 69% imitation of the 'head-action' in our replication of Meltzoff (1988), only 7% of the 14-month-olds imitated the very same head-bending action in this 'no-effect' condition. This suggests that (a) the ostensive-communicative cues are in themselves not sufficient to trigger imitation, and (b) the fact that the changed context of the same behavior probably rendered it functionally uninterpretable for these young infants, resulted in the disappearance of its imitation.

So maybe Tomasello's full 'cognitive transparency criterion' should be relaxed into some more general 'schematic cognitive interpretability requirement'. On this account, the behavior should receive an at least partially completed, even if "conceptually shallow or schematic" interpretation (Keil, 2003) in terms of one of the core interpretive "modes of construal" (Keil, 1995; Kelemen, 1999a, b; Gergely & Csibra, 2003) that infants have at their disposal (such as their teleo-functional means-end scheme, causal-physical scheme of contact and force dynamics, or their understanding of distal referential relations as exemplified by eye-gaze, pointing or naming behaviors). In fact, it is in relation to such a schematic and only partially completed cognitive functional interpretation (e.g., that the head-touch behavior functions as a means to a goal) that the particular choice of the behavior manifested as the means remains cognitively 'opaque' to the infant. If such an apparently unjustifiable behavioral choice is, nevertheless, ostensively manifested in a pedagogical context, infants

will interpret this as conveying significant and relevant cultural knowledge that is new for them and so should be retained and imitatively reproduced (even though it may remain cognitively 'opaque' to them).

This 'pedagogically guided social learning strategy' may be seen as a developmental example of what Keil (2003) refers to as "the benefits of being [conceptually] shallow" (p. 372). He points out that "people…rapidly decide which domain of causal patterns is relevant and then use their own schematic knowledge of relations and patterns to constrain explanations on the fly.…Adults and children alike amplify their understandings by relying on the division of cognitive labour that is intrinsic to all cultures.…One advantage of lean causal representations — Keil emphasizes — may be rapid development" (pp. 371–2).

Recently, we also ran a further control study that was identical to the original 'Hands-free' condition, except for the fact that — following the ostensive-communicative cues — the demonstrator manifested *both* the (unusual) 'head-action' and the (predictable) 'hand-action'. Again, we found that imitation of the 'head-action' has practically disappeared as a function of this contextual change: only one of 14 fourteen-month-olds imitated the less efficient means act demonstrated (the 'head-action'), while all subjects performed the predictable 'hand-action'. It seems, therefore, that the ostensive manifestation of both the head- and hand-actions 'sanctioned' both actions as culturally equally acceptable and relevant alternative means to the goal. Since in this way the pedagogically transmitted information did not compete with considerations of physical efficiency, the infants' choice of behavior was fully determined by the latter, and no imitative learning of the less effective 'head-action' took place.

The potency of the ostensive-communicative demonstration context in socially inducing fast learning of the relevant function of new artifacts has also been elegantly demonstrated by Deb Kelemen and her colleagues' recent studies on the social determinants of the early understanding of artifact functions (DiYanni & Kelemen, 2005; Casler & Kelemen, 2005). For example, Casler and Kelemen (2005) have shown that 2.5-year-old children rapidly form a teleofunctional representation of a novel instrumental tool after only one single ostensive demonstration. In their studies, children were presented with two tools that were physically equally affordant for a new task (turning on a light box). The children were first allowed to explore the relevant physical affordance properties of the two tools (they put both into slots), but then they saw only one being chosen and demonstrated to perform the new function. At 2.5 years of age, children repeatedly returned to the demonstrated tool as "for" the task, both immediately and after a multi-day delay, despite the ready availability of

the alternative. Children were also found to dissociate, preferring to use the alternative when asked to perform a different function (crushing up crackers). Casler and Kelemen (2005) argue that after only one single demonstration indicating the new artifact's functional use in a pedagogical context, "children will construe the tool as for that particular purpose and…avoid using it for another feasible purpose". Furthermore, the two-year-olds appeared to view the demonstrated function as an intrinsic property of the object that should be readily recognizable by others as well. This is shown by the fact that they also expected *another person,* who was absent during the demonstration and was unfamiliar with the two new tools, to choose to use for the new function the same tool that the children had seen contrastively chosen by the demonstrator earlier. This provides support for our hypotheses that (a) pedagogical cues trigger fast learning, and (b) the infant's 'pedagogical stance' contains a "universality assumption" that the cultural contents conveyed through ostensive-communicative manifestations constitute publicly shared and generalizable knowledge.

But is it really the case that the kind of inferences and interpretations underlying the selective nature of imitative learning are triggered only if the observed target action is manifested in a pedagogical context? To find out we have recently run a new version of the Gergely et al. (2002) study (Király, Csibra, & Gergely, 2004). Half of the subjects were presented with the 'head-action' in either the 'Hands-free' or the 'Hands-occupied' context-conditions demonstrated with rich ostensive-communicative cues as before. The rest of the 14-month-olds participated in an "incidental observation" situation in which they observed the very same 'head-action' in either the 'Hands-free' or the 'Hands-occupied' context-condition, but without being exposed to any ostensive-communicative cues by the model. Our findings indicate that the pedagogical demonstration context does make a qualitative difference. In the pedagogical demonstration situation we have replicated (now the third time) the same pattern of selective imitation of the 'head-action' as in Gergely et al. (2002). Furthermore, the 'head-action' was imitated in the 'Hands-free' condition significantly more when preceded by ostensive-communicative cues than when only incidentally observed in a non-communicative context. In fact, while we did find some imitation of the 'head-action' in both of the two 'incidental observation' conditions as well, the selective degree of imitation present in the pedagogical cuing condition has disappeared: there was no differential imitation evoked in the "Hands-free" *versus* "Hands-occupied" context-conditions when no ostensive-communicative cues were present.

Conclusions

We reviewed recent evidence revealing (a) the selective interpretive nature of imitative learning in human infants and (b) the role of ostensive-communicative cues in constraining and guiding the infant's selective interpretation of what is the new and relevant cultural information conveyed by the other's manifestation that should be fast learned. We argued that these findings pose problems for current theories of human imitative learning whose predictive scope is either too broad or too narrow to account for the type of selectivity that characterizes imitative learning in infants. We proposed a new theory of human cultural learning in which imitation is seen as a basic mechanism that has been recruited, guided and constrained by the human-specific adaptation for 'pedagogy', a complex cognitive system of mutual design that is dedicated to the fast and efficient transmission of cultural knowledge in humans. We argued that the selective interpretive nature of early imitative learning can be explained as a result of the implicit assumptions built into the infant's 'pedagogical stance' that constrain and guide imitative learning, and that is activated by the ostensive-communicative cues of knowledgeable others who manifest new and relevant cultural information for the infant to learn.

Acknowledgment

This research was supported in part by a Fellowship of the J. S. Guggenheim Memorial Foundation to the first author, and by grants from the Hungarian National Science Foundation (OTKA #T034567) and The Leverhulme Trust.

Notes

1. For a critical analysis of this position, see Gergely & Csibra (2005b).

2. For example, after repeatedly seeing an actor approach its goal by jumping over an obstacle, infants show surprise (look longer) when — following the removal of the obstacle — the actor performs the previous jumping action again to get to the goal (this time, however, jumping over nothing). In contrast, when the actor *changes* his behavior in a justifiable manner approaching the goal through the most direct straight-line path that has become available (rational goal-approach), the infants look significantly less (showing no sign of surprise) at this novel (but sensible) action (Gergely et al., 1995; Gergely & Csibra, 2003).

3. The ostensive *manifestation* of a motor skill involves a saliently transformed manner of motor execution when compared to its primary functional *use*. Think of the difference between hammering a nail in *vs.* demonstrating to a novice *how* to hammer a nail in. Manifestations involve slowed-down, schematic, exaggerated, or sometimes only partially executed transformations of the primary motor program that foregrounds and thus helps to identify the relevant and new information for the novice to acquire.

4. See Csibra and Gergely (2005) for arguments showing that many early emerging social cognitive capacities — such as social referencing (Egyed, Király, & Gergely, 2004), protodeclarative pointing, or word learning — can be usefully reinterpreted as examples of cultural learning through pedagogy.

5. Note that these assumptions are directly analogous, if not identical, to the Gricean pragmatic assumptions of ostensive communication as spelled out in Sperber and Wilson's (1986) relevance theory. In our view, however, pedagogy is a primary adaptation for cultural learning and not a specialized module dedicated to the recovery of speaker's intent in linguistic communication that has evolved later as a sub-module of human theory of mind (Sperber and Wilson, 2002).

References

Byrne, R. W., & Tomasello, M. (1995). Do rats ape? *Animal Behaviour, 50,* 1417–1420.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competences from 9 to 15 months of age. *Monographs of the Society of Research in Child Development, 63(4).*

Casler, K., & Kelemen, D. (2005). Young children's rapid learning about artifacts. *Developmental Science* (in press).

Csibra G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science, 1*:2, 255–259.

Csibra, G., & Gergely, G. (2005). Social learning and social cognition: The case of pedagogy. In M. H. Johnson & Y. Munakata (Eds.), *Progress of Change in Brain and Cognitive Development. Attention and Performance XXI.* Oxford: Oxford University Press.

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science, vol. 27*(1), 111–133.

DiYanni, C., & Kelemen, D. (2005). Using a bad tool with good intention: How preschoolers weigh physical and intentional cues when learning about artifacts. *Manuscript under revision*.

Egyed, K., Király, & Gergely, G. (2004). Object-centered versus agent-centered Interpretations of referential attitude expressions in 14-month-olds. Poster presented at the *14th Biennial International Conference on Infant Studies*, May 2004, Chicago, IL, USA.

Gergely, G., & Csibra, G. (2003). Teleological reasoning about actions: The naïve theory of rational action. *Trends in Cognitive Sciences, 7*, 287–292.

Gergely, G., & Csibra, G. (2005a). Sylvia's recipe: Human culture, imitation, and pedagogy. In S. Levenson & N. Enfield (Eds.), *Roots of Human Sociality: Culture, Cognition, and Human Interaction.* Oxford: Berg Publishers, (in press).

Gergely, G., & Csibra, G. (2005b). A few reasons why we don't share Tomasello et al.'s intuitions about sharing. A commentary on Tomasello et al.'s "Understanding and sharing intentions: The origins of cultural cognition." *Behavioral and Brain Sciences* (in press).

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature, Vol. 415*, p. 755.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, Vol. 56*, No. 2., 165–193.

Goodall, J. (1986). *The chimpanzees of Gombe.* Cambridge, Mass.: Harvard University Press.

Heyes, C. M. (1993). Imitation, culture and cognition. *Animal Behaviour, 46,* 999–1010.

Heyes, C. M., & Galef, B. G. (1996). *Social learning in animals: The roots of culture*. NY: Academic Press.

Keil, F. (1995). The growth of understandings of natural kinds. In D. Sperber, D. Premack, & A. Premack (Eds.), *Causal cognition* (pp.234–267). Clarendon Press.

Keil, F. (2003). Folkscience: coarse interpretations of a complex reality. *Trends in Cognitive Sciences, 7*, 368–373.

Kelemen, D. (1999a). Function, goals and intention: children's teleological reasoning about objects. *Trends in Cognitive Sciences, 12*, 461 — 468.

Kelemen, D. (1999b). The scope of teleological thinking in preschool children. *Cognition, 70,* 241–272.

Király, I., Csibra, G., & Gergely, G. (2004). The role of communicative-referential cues in observational learning during the second year. Poster presented at the 14th Biennial International Conference on Infant Studies, May 2004, Chicago, IL, USA.

Meltzoff, A. N. (1988). Infant imitation after a one week delay: Long term memory for novel acts and multiple stimuli. *Developmental Psychology, 24*, 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 C. M. Heyes & B. G. Galef (Eds), *Social learning in animals: The roots of culture* (pp. 347–370). NY: Academic Press.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science, 198*, 75–8.

Meltzoff, A. N., & Moore, M. K. (1989). Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms. *Developmental Psychology, 25,* 954–62.

Meltzoff, A. N., & Moore, M. K. (1997). Explaining facial imitation: theoretical model. *Early Development and Parenting, 6,* 179–92.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell.

Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind & Language, 17(1),* 3–23.

Thorpe, W. H. (1963). *Learning and instincts in animals*. London: Methuen.

Tomasello, M. (1996). Do apes ape? In C. M. Heyes & B. G. Galef (Eds), *Social learning in animals: The roots of culture*. NY: Academic Press.
Tomasello, M. (1999). *The cultural origins of human cognition.* Boston: Harvard University Press.
Tomasello, M., & Call, J. (1997). *Primate cognition*. Oxford: Oxford University Press.
Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences, 16,* 495–552.
Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* (in press).
Want, S. C., & Harris, P. L. (2002). How do children ape? Applying concepts from the study of non-human primates to the developmental study of 'imitation' in children. *Developmental Science, 5(1),* 1–13.
Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C. E. G., Wrangham, R. W., & Boesch, C. (1999). Cultures in chimpanzees. *Nature*, *399,* 682–685.

Authors' addresses

György Gergely
Head of Department of Developmental Research
Institute for Psychological Research
Hungarian Academy of Sciences
Victor Hugo u. 18–22
1132 Budapest
HUNGARY
E-mail: gergelyg@mtapi.hu

Gergely Csibra
Centre for Brain and Cognitive Development
School of Psychology
Birkbeck College
Malet Street
London WC1E 7HX
UK
E-mail: g.csibra@bbk.ac.uk

About the authors

György Gergely is a professor at the Institute for Psychological Research of the Hungarian Academy of Sciences, where he is the Head of the Department of Developmental Research. He received his B.Sc. from The London School of Economics in 1975, his M.Phil. from University College London in 1978, and his Ph.D. from Columbia University, N.Y., in 1986. He was a postdoctoral Fellow at Stanford University. His research interests include early cognitive and socio-emotional development in infancy, theory of mind, attachment and developmental psychopathology.

Gergely Csibra is a senior researcher at the Centre for Brain and Cognitive Development, Birkbeck College, London, and a senior lecturer at the School of Psychology at Birkbeck College, London. He received his Ph.D. in 1992 at ELTE University, Budapest. His research interests include developmental cognitive neuroscience, infant cognitive development, and theory of mind.