The Rationality of (Over)imitation

Stefanie Keupp1,2, Tanya Behne2,3, and Hannes Rakoczy2,3
1Cognitive Ethology Lab, German Primate Center, Göttingen, Germany; 2Leibniz-ScienceCampus Primate Cognition, University of Göttingen; and 3Institute of Psychology, University of Göttingen

Abstract
Imitation is a powerful and ubiquitous social learning strategy, fundamental for the development of individual skills and cultural traditions. Recent research on the cognitive foundations and development of imitation, though, presents a surprising picture: Although even infants imitate in selective, efficient, and rational ways, children and adults engage in overimitation. Rather than imitating selectively and efficiently, they sometimes faithfully reproduce causally irrelevant actions as much as relevant ones. In this article, we suggest a new perspective on this phenomenon by integrating established findings on children’s more general capacities for rational action parsing with newer findings on overimitation. We suggest that overimitation is a consequence of children’s growing capacities to understand causal and social constraints in relation to goals and that it rests on the human capacity to represent observed actions simultaneously on different levels of goal hierarchies.

Keywords
overimitation, rational imitation, rationality, goals, action parsing

Imitation plays a fundamental role in people’s social life and may even be a uniquely human capacity (Tennie, Call, & Tomasello, 2009). Ontogenetically, imitation emerges very early in infancy (Meltzoff & Moore, 1989). However, recent research on children’s and adults’ action imitation has yielded a surprising picture: On the one hand, infants as young as 1 year are able to engage in selective and rational imitation. That is, when reproducing an experimenter’s goal-directed behavior, they choose the most efficient means to attain the goal and safely ignore any superfluous or causally irrelevant actions that were included in the demonstration (Brugger, Lariviere, Mumme, & Bushnell, 2007). On the other hand, children (from around age 2 or 3) and adults frequently imitate in seemingly unselective and irrational ways, faithfully imitating any kind of behavior, including bizarre and causally irrelevant actions, sometimes even to the extent of jeopardizing goal attainment. A classic way to study imitation strategies is to present participants with a puzzle box (e.g., Horner & Whiten, 2005; Lyons, Damrosch, Lin, Macris, & Keil, 2011). In this paradigm, an experimenter demonstrates a number of actions on a box that has multiple movable parts and contains a reward. Some of the actions are necessary to obtain the reward (causally relevant actions; e.g., the door hiding the reward must be opened), whereas others are completely irrelevant for reward retrieval in obvious and perceivable ways (causally irrelevant actions; e.g., tapping on the box and moving attachments on the outside of the box; see Fig. 1 for a schematic depiction of a typical procedure). In this task, children younger than 2 years frequently produce only goal-relevant actions and ignore the causally irrelevant ones, but older children include both types of actions when it is their turn to operate the box (e.g., McGuigan & Whiten, 2009).

Numerous studies have documented this phenomenon that has become known as overimitation (i.e., the reproduction of causally irrelevant action elements within a goal-directed action sequence). Specifically, when a demonstrator intentionally performs a goal-directed action sequence that includes an element that is obviously causally irrelevant, observers’ reenactments that include the causally irrelevant element are referred to as cases of overimitation. Such inefficient and
The Rationality of (Over)imitation

Fig. 1. The typical procedure of an overimitation study. In the demonstration phase, the model performs the causally irrelevant action, A, and then the causally relevant action, B. For example, as shown here, the model may tap on a box with a stick (action A) and then open the door of the box (action B), so that the reward hidden in the box becomes available (effect E). In the imitation phase, it is the child’s turn. A common finding is that a majority of children perform both the causally irrelevant action and the causally relevant action. Finally, in the third-party observation phase, another agent, often a puppet, gets a turn. Sometimes this agent performs both actions (illustration on the left), and sometimes he or she performs only the causally relevant action (illustration on the right). Children frequently spontaneously criticize the agent for making a mistake when the agent omits the causally irrelevant action. This indicates that they interpret the demonstrated action chain not merely in terms of an activity leading to a functional outcome (getting the reward) but also as a generic means of how things should be done by everyone who engages in the demonstrated activity.
unselective forms of imitation stand in stark contrast to the rational, selective imitation shown already by infants. These findings present a puzzling phenomenon: Older children jeopardize successful or efficient task performance, but infants select instrumentally efficient means.

Different accounts regarding the psychological mechanisms underlying this phenomenon have been put forward. Some focus on the important role of social motivations: By copying exactly, an imitator aims at establishing a social bond, signaling group membership or sameness; that is, the imitator pursues affiliative motives toward the model (e.g., Nielsen & Blank, 2011; Over & Carpenter, 2012). Normative accounts suggest that more generic normative demands motivate the imitator's behavior (e.g., Kenward, 2012; Keupp, Behne, & Rakoczy, 2013). One consequence of the normative view is that the imitator will assume that not only the imitator and model but everyone ought to adhere to such normative demands, and failing to do so constitutes a mistake. A third class of accounts draws on the imitator's causal understanding. One of the most prominent is the automatic-causal-encoding account, which suggests that the imitator falsely assumes all actions to be causally relevant for reaching the goal after having seen a model performing them intentionally (e.g., Lyons, Young, & Keil, 2007). Another account proposes that all actions are assumed to be somehow relevant and thus are copied at first, but that there is the possibility for later refinement (e.g., Flynn, 2008; Horner & Whiten, 2005). There is now growing consensus that overimitation is a multifaceted phenomenon and that none of the suggested underlying psychological mechanisms alone can account for all instances of overimitation. But researchers continue to treat overimitation as a puzzling, specific phenomenon in need of explanation. In fact, a number of publications have referred to overimitation as "mysterious" and "curious" (Lyons et al., 2011, p. 1158), as "puzzling" (Kenward, Karlsson, & Persson, 2011, p. 1239), and as "mindless or blinkered over-copying" that is distinguished from the "sophisticated and rational selectivity" of selective imitation (Whiten, McGuigan, Marshall-Pescini, & Hopper, 2009, p. 2425).

**Different Forms of Imitation Reflect Different Forms of Action Interpretation**

Here we present a new argument and review recent evidence in its favor in an effort to address this issue and resolve the apparent conundrum regarding the seemingly unselective and inefficient ways older children and adults, but not infants, imitate. We suggest that there is in fact nothing particular or peculiar about overimitation that is in need of specific explanation. Key to our suggestion is the insight that rational and apparently irrational forms of imitation are merely different manifestations of the human capacity for flexible action interpretation, action selection, and action execution (Schachner & Carey, 2013). In particular, some actions are interpreted as purely instrumental; that is, the elements in a behavioral sequence are considered mere causal means to ends. Other actions, in contrast, are interpreted as conventional activities in which the behavioral elements are essential conventional (rather than causal) parts of bigger action chains. In the former case, it might be sufficient to reproduce only some of the actions or even achieve the goal by a different functional means, whereas in the latter case, the (perceived) conventionality of the activity dictates faithful reproduction of all the actions. For example, moving an object to a specific location by kicking it with your foot is a rather inefficient choice of means when the task is to tidy up. However, the very same behavior is an essential, even constitutive, part of the activity—and one you may become world-famous for—if performed in the context of a football game. The way agents imitate will thus be a function of their specific interpretation of the action they saw and their perception of the requirements for action execution in the given situation.

Such flexible action parsing and imitation has been documented in infants, children, and adults, who, for instance, imitate actions as a function of the salience of a goal (e.g., Bekkering, Wohlschlager, & Gattis, 2000; Carpenter, Call, & Tomasello, 2005; Schachner & Carey, 2013; Southgate, Chevallier, & Csibra, 2009; Williamson & Markman, 2006; Wohlschlager, Gattis, & Bekkering, 2003). For example, when adults and children were asked to imitate an experimenter who moved his hand and touched a table, their imitation was more precise and faithful (they tended to copy the experimenter’s choice of hand) when the table was unmarked than when it was marked with two dots (Bekkering et al., 2000; Wohlschlager et al., 2003). These results indicate that participants imitate actions more or less faithfully depending on what they understand the most important goal of the task to be: If there is an instrumental focus, such as an external goal to be attained ("touch this dot"), both young children and adults fulfill the goal but do not necessarily follow the exact means demonstrated. However, if the very same behavior is demonstrated in the absence of an external goal, thus suggesting a conventional focus ("move your arms in such and such ways"), both young children and adults do follow the exact means demonstrated. This interpretation is in line with recent findings from Schachner and Carey (2013), who showed that adults infer
movement itself as the goal when visual goals are absent or when intentional movements are inefficient means of achieving the external goal.

This general principle, we argue, also applies to what has come to be known as overimitation: The perceived requirements for action execution depend on which goal is perceived as the hierarchically most important one in a given situation. Thus, children observe a behavior, interpret it against the background of contextual cues (which can involve, for instance, evaluation of the conventional or instrumental relevance of available action representations given the context), and execute their own response depending on which action interpretation they choose as the most appropriate (which also depends on their own goals and circumstances).

The question of why humans overimitate can be addressed on two levels: First, the ultimate level concerns the evolutionary function of overimitation, and second, the proximate level concerns the cognitive mechanism (or mechanisms) that cause such imitation patterns. In this article, we are mainly concerned with the underlying cognitive processes.

**Overimitation Is More Flexible and Sophisticated Than Previously Assumed**

Initial research on overimitation indicated that it is an inefficient imitation strategy that occurs robustly across a wide range of tasks and domains and that is characterized by a lack of context sensitivity, flexibility, and rationality (see, e.g., the automatic-causal-encoding strategy suggested by Lyons et al., 2011). However, upon closer inspection, overimitation turns out to be a sophisticated, flexible, and context-sensitive behavior that is intimately connected to rational interpretation of actions. In fact, recent studies demonstrated that children’s overimitation is context-sensitive in the following three ways.

First, children’s responses are influenced by the context in which an action is demonstrated. Depending on the general setting, the situational constraints, and the way the action is communicatively introduced, children may interpret the same token of behavior in different ways. For example, they may regard it as accidental, intentional, or even pedagogical and reenact it accordingly (Butler & Markman, 2012; Carpenter, Akhtar, & Tomasello, 1998; Meltzoff, 1995). For the current argument, and because it is probably the dimension most relevant to overimitation, we focus on the distinction between an instrumental and a conventional stance. An instrumental stance will lead to the selective and efficient imitation of merely those means that seem causally necessary to reproduce the desired ends (Williamson, Meltzoff, & Markman, 2008). In contrast, a conventional stance on the very same token of behavior manifests itself naturally in faithful overimitation (Clegg & Legare, 2016; Kenward, 2012; Keupp et al., 2013; Legare, Wen, Herrmann, & Whitehouse, 2015; Yu & Kushnir, 2014).

Second, children’s responses are influenced by the context of action execution. When children observe a demonstration, they can adjust their subsequent action execution and responses to third parties’ behavior flexibly as a function of task requirements or social context. If the task or context requires conventional precision, they will overimitate; if, in contrast, the task or context require instrumental efficiency, they will refrain from overimitation (Keupp, Behne, Zachow, Kasbohm, & Rakoczy, 2015; McGuigan & Robertson, 2015; Moraru, Gomez, & McGuigan, 2016; Nielsen & Blank, 2011). These findings indicate that children have different representations of the observed behavior available and choose to enact the particular course of action that they deem most appropriate given their circumstances: For example, they overimitate more often in conventional than in instrumental contexts and more often in public contexts involving peer pressure than in private contexts (see Fig. 2).

Third, children’s responses are influenced by the consequences that these actions entail. Children often overimitate in the absence of any competing reasons not to do so but can shift strategies toward efficient imitation in contexts in which overimitation has costly consequences. A recent study implemented an overimitation scenario in which irrelevant actions sometimes had morally costly consequences (i.e., the actions resulted in the destruction of a valuable item that belonged to another agent; see Fig. 2). Children often overimitated a superfluous action when it had no relevant consequences. But they did so significantly less often when the superfluous action was coupled with costly consequences. Note that they still imitated the costly actions more in the conventional than the instrumental context, which indicates that they parsed the demonstration rationally, giving due consideration to conventional, instrumental, and moral demands (Keupp, Banckcn, Schillmöller, Rakoczy, & Behne, 2016).

The findings we have just described have been corroborated by additional measures involving third-person consequences: When children interpret an action sequence as a token of a conventional action type, they not only act accordingly by imitating the sequence faithfully, but also expect third parties to follow the conventions as well. Omission of causally irrelevant actions is rational, and thus acceptable, when the goal is to bring about an instrumental outcome (hence, no normative intervention is required). But the same
Fig. 2. Schematic depiction of various forms of context-specific imitation and overimitation. In the demonstration phase, the model performs the causally irrelevant action, A, and then the causally relevant action, B, and achieves effect E. In the imitation phase, children tend to overimitate (illustrations on the left) if they are given instructions that highlight the conventional nature of the task, if they are witnessed by other people who have modeled the action sequence A + B, or if overimitation does not involve high costs. In contrast, children tend to refrain from overimitation (illustrations on the right) if they are given instructions that highlight the instrumental nature of the task, if they are unobserved, or if overimitation would involve high costs. Note that these schematic depictions are for illustration purposes only and do not necessarily represent the procedures in the original studies.
omission constitutes a mistake when the goal is to perform a conventional activity of which the causally irrelevant act is a constitutive part. Thus, enforcing imitation of causally irrelevant actions by third parties is rational considering the generic nature of conventions. In line with this argument, studies have shown that in the conventional, but not the instrumental, context, children systematically criticize third parties who imitate efficiently and omit action elements that are causally irrelevant, yet considered to be conventionally essential (Kenward, 2012; Keupp et al., 2013). In addition, children integrate different conflicting demands and not only reduce their own overimitation in a conventional context when consequences are costly, but also reduce their criticism of third parties who omit the irrelevant action in these conditions (Keupp et al., 2016).

**Integrating Goals, Action Parsing, and Imitation**

Taken together, these findings shed novel light on overimitation: Overimitation is not so surprising after all, but rather is a particular expression of general and flexible capacities for interpretation, selection, and execution of actions. In fact, the findings suggest that children's imitative behavior may occur at the "program level" (Byrne & Russon, 1998) and that selective overimitation reflects flexible parsing of the actions in an action hierarchy as a function of the overarching goal. The view of overimitation that we put forward integrates research on overimitation with research on the development of imitation more generally—a field that has long been studying how infants and young children interpret actions in terms of hierarchical goals and how sophisticated, flexible, and context-sensitive their imitative responses are.

Strictly speaking, overimitation is a misnomer. It conveys the misconception that there is something wrong with the behavior—and that a separate explanation for this phenomenon is needed. This implication goes back to the pioneering researchers who coined the term against a background of rationality defined narrowly in terms of instrumental efficiency (Lyons et al., 2007). However, if one considers the background of human cultural ecology, it is highly misleading to restrict rationality to instrumental efficiency. Actions that would be pointless from a purely instrumental perspective often are meaningful from a conventional perspective. Just think, for example, of fine-grained movements in dance or sport that serve no further end beyond their own execution. This idea has long been captured in more differentiated and nuanced conceptions of the varieties of rationality, such as Max Weber's (1922/1978) famous distinction between “instrumental rationality” (actions as mere means to ends) and “value rationality” (actions as ends in themselves). It is also reflected in current distinctions between instrumental and ritual stances in action interpretation (Clegg & Legare, 2016) and between the instrumental and social functions of action imitation (Over & Carpenter, 2013; Uzgiris, 1981). Different criteria of rationality thus underlie humans’ action interpretation: These criteria include both standards of instrumental efficiency and standards of social and conventional norms that define the ways in which actions, including causally irrelevant ones, make conventional sense. And these different (instrumental vs. conventional) standards result in quite different, yet equally rational, forms of selective or faithful imitation.

Imitation can thus be properly understood only in its broader relations to naive folk theories of action and rationality (Gergely & Csibra, 2003; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016). One of the fundamental characteristics of everyday action understanding is that actions are perceived, parsed, individuated, described, and interpreted hierarchically in different and flexible ways (formally, this notion can be captured by action trees; see, e.g., Goldman, 1970). One and the same behavior (say, pressing a switch on an electric kettle) can be described at different levels of specificity and hierarchy (e.g., moving a thumb, pressing a switch, switching on the kettle, boiling water, making tea, making breakfast). This feature of action understanding is known as the accordion effect (Bratman, 2006; Feinberg, 1970). What is crucial about this for present purposes is that the relations between a given action and the action elements that it comprises can be of different logical kinds. Some action elements are causally necessary to implement an instrumental action one level up. That is, the higher-level action is performed by means of the lower-level ones. For example, boiling water is an instrumental action: It is fulfilled when the water boils and can be causally realized by various means (e.g., by switching on the kettle). In contrast, some action elements are not causally linked to actions one level up, but rather are conventionally and constitutively linked to them. That is, the higher-level action is performed by virtue of performing the lower-level ones. Scoring a goal in football (soccer) is not a purely instrumental action that is fulfilled when the ball is in the net and thus can be causally realized in various ways (e.g., by throwing the ball into the net or by using a car to drive the ball into the net). Rather, certain means (e.g., using the foot or head to send the ball into the net) are conventionally essential elements of scoring a goal in football.

Thus, different forms of imitation rest on different forms of action interpretation and their corresponding rationality criteria: A given behavior can be perceived
as purely instrumental, comprising subordinate elements that are related to it in causal ways. In imitative reproduction, the aim is thus to reproduce the topmost instrumental action in causally efficient ways under the given circumstances, and the relevant rationality criteria are those of efficiency. Some (causally irrelevant) elements may be omitted, and the actions used to achieve the goal may be more efficient alternatives to the originally demonstrated ones. The very same behavior, however, can—under other circumstances—be perceived as a conventional action comprising some causally irrelevant yet constitutive elements that are necessary from a conventional point of view. In imitative reproduction, the aim is thus to reproduce the topmost conventional action by following its constitutive (rather than causal) elements, and the corresponding rationality standards are those of conventional faithfulness and accuracy. People can thus form different representations of an observed behavior and, if asked to act themselves, produce the behavior that seems appropriate given their own context-dependent action interpretation. In addition, the realized behavioral response will depend on the observers’ own goals; for instance, they may aim to comply with a conventional norm, to be efficient, or to use imitation as a means of bonding with the model (Hilbrink, Sakkalou, Ellis-Davies, Fowler, & Gattis, 2013; Nielsen, 2006; Over & Carpenter, 2009).

From a developmental point of view, different forms of imitation emerge as a function of the interplay of social-cognitive capacities. From infancy on, children operate with a basic naive theory of action and rationality: Numerous studies have shown that when infants observe and interpret actions, they apply a principle of rationality which presupposes that people (or other agents) perform the most efficient actions that fulfill their goals given the current physical environment (Gergely & Csibra, 2005). Capacities for flexibly parsing actions in intentionally meaningful units have been found in infant as young as 10 months old (Baldwin & Baird, 2001; Baldwin, Baird, Saylor, & Clark, 2001), and infants around this age begin to understand that goals are independent of actions and that some actions—specifically, those causally related to the goal—are performed as means of attaining a goal (Sommerville & Woodward, 2005a, 2005b; Woodward & Sommerville, 2000). And from the age of 2 to 3 years, children are sensitive to the conventionality and normative structure of social actions (Diesendruck & Markson, 2011; Rakoczy & Schmidt, 2013). Thus, beyond their understanding of the physical causal constraints, they begin to understand that actions can also be socially constrained. Over development, and in ways yet to be understood in detail, children integrate these capacities (see, e.g., Buchsbaum, Gopnik, Griffiths, & Shafto, 2011) and put them to use in a more comprehensive and systematic framework for parsing, interpreting, and imitating actions that incorporates both instrumental and conventional notions of rationality. Sophisticated understanding of different types of constraints broadens the scope of possible representations and action goals that can be derived from an observed behavior. Thus, infants can infer intended goals from (circumstantial) demonstrations and act in rationally efficient ways by imitating merely the outcome-related actions. Older children also understand otherwise meaningful relations and might choose to include causally irrelevant elements that they understand as socially or normatively constitutive parts of the actions—which results in faithful overimitation in some cases.

**Concluding Remarks and Future Directions**

Imitation itself is thus a very sophisticated and coherent phenomenon. Different forms of imitation—be it efficient imitation or faithful imitation—are equally rational, just in different respects and in light of different (instrumental vs. conventional) standards of rationality. The research reviewed here suggests that in their interpretation, selection, and execution of actions, children and adults operate with varying stances that embody these different rationality standards.

However, little is known so far about when and why each of these stances comes to the forefront. Some recent research, for example, has begun to uncover conditions under which children and adults tend to interpret actions as tokens of generic, conventional types (Csibra & Gergely, 2009). For example, studies have demonstrated that overimitation is more likely if the demonstrator is knowledgeable rather than naive (Buchsbaum et al., 2011) or belongs to the observer’s minimal group rather than the opposing one (Gruber, Descheniaux, Frick, & Clément, 2017). Overimitation is also more likely if the demonstration of the causally inefficient strategy precedes, rather than follows, the demonstration of a more efficient strategy (Schleihauf, Graetz, Pauen, & Hoehl, 2018), and if the children tested are typically developing rather than on the autism spectrum (Vivanti, Hocking, Fanning, & Dissanayake, 2017; but see Nielsen, Slaughter, & Dissanayake, 2013). Recent research has also begun to explore the extent to which children’s learning through imitation is facilitated by or reliant on child-directed demonstrations and teaching (for a review and discussion, see Shneidman & Woodward, 2016), and the conditions under which such demonstrations limit children’s own spontaneous discovery and innovation (Bonawitz et al., 2011; Carr, Kendal, & Flynn, 2015).
Questions to be explored in future research include the following: Which communicative aspects of a demonstration context evoke the different kinds of stances and interpretations and thus influence the course of action that will be considered most rational? Are there interesting age differences and developmental patterns in the ways that action demonstration and communication evoke the different stances, and, if so, what may be the crucial social-cognitive correlates and underpinnings of relevant developmental shifts? Another open question is how the different stances and standards that underlie people’s naive theory of action and rationality are cognitively and computationally realized. Exciting new modeling work suggests the possibility, for example, that the fundamental structure of naive action theory may be formalized in a broadly Bayesian framework (Jara-Ettinger et al., 2016) and may open the way for developing more precise and formal accounts of the varieties of rational imitation and its cognitive foundations.

We have focused on proximate questions about the cognitive foundations and development of different forms of imitation. However, from an evolutionary point of view, imitation is a powerful social learning strategy that poses many intriguing ultimate questions. Faithful imitation, including overimitation, is important for promoting cumulative culture by supporting the transmission of instrumental skills and opaque cultural knowledge, such as group-specific traditions (Flynn & Smith, 2012; Legare & Nielsen, 2015; Nielsen, Mushin, Tomaselli, & Whiton, 2014). Selective imitation, on the other hand, is necessary for innovation and improvement of established techniques and practices. How might flexible, rational use of selective and faithful imitation strategies have developed? One broad possibility is that the various forms of imitation constitute distinct adaptations (and that the need to coordinate them flexibly perhaps even presented an evolutionary motor for systematic action interpretation). An alternative possibility, however, is that these different forms of imitation are a by-product of humans’ more general capacities for observing, parsing, interpreting, and executing actions. From a comparative perspective, one may ask whether anything remotely like overimitation is present in nonhuman animals at all (for negative findings in great apes and canids, see Clay & Tennie, 2017; Horner & Whiton, 2005; Johnston, Holden, & Santos, 2017; Nielsen & Susianto, 2010).

Finally, does our analysis suggest that all forms of imitation will turn out to be rational in the end? Considering imitation broadly as encompassing any form of reproduction of observed behavior, there surely remain many forms of imitation—ranging from compulsive imitation in neuropsychological patients to unconscious mimicry—that resist any reconstruction as rational. A big challenge for future research is thus to delineate more clearly the dividing line between those forms of imitation that are rational at first sight, those that turn out to be rational after closer examination, and those that remain nonrational even upon such inspection.

**Action Editor**

Marjorie Rhodes served as action editor and June Gruber served as interim editor-in-chief for this article.

**ORCID iD**

Stefanie Keupp https://orcid.org/0000-0002-5451-4256

**Acknowledgments**

We thank Daniela Fuchs for her help with preparing the figures.

**Declaration of Conflicting Interests**

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

**References**


Keupp et al.


