

Chapter 8

From Parsing Actions to Understanding Intentions

Richard W. Byrne

Abstract “Having a theory of mind” is often invoked to explain remarkable abilities in social cognition, but in reality this is little more than a re-description of the data, a challenge for theorists to understand what it really means, and how we—and perhaps some other species—evolved those abilities. I argue that these abilities most likely grew out of an understanding of action, and in particular that the key first step was the evolution, by shared ancestors of modern humans and the living great apes, of an ability to parse the skilled actions of others. Parsing organized, goal-directed behaviour allows the intention of the model to be “seen” (as the typical result which terminates action), and the cause-effect relations in the steps of the process to be “seen” (as the sequence and coordination of actions necessary for that result to be achieved). In contrast, reasoning about true intentionality and true causality may depend on possession of language, and be of relatively minor usefulness in everyday interactions.

When we notice someone engaged in activity, we see not only how their body moves and what effects those movements are having on other things, but we also see what it *means*. The meaning of action includes *what is likely to happen next*, as a consequence of what has been done already; and what overall result is to be expected from the activity, in short, *why it is being done*. This description applies to the simplest of organized, purposeful actions but also to what is arguably our most sophisticated cognitive ability, the ability to talk. When we hear someone talking our language, we don’t merely register a series of sounds, phonemes, words, phrases, and meanings; we immediately have some understanding of what thoughts have led to their speaking, whereabouts (metaphorically) their speech is going, and what pragmatic effects the speaker might be trying to achieve by it. These observations are so familiar and commonplace that normally we pay them no heed: rather, we only notice when people do things that make no sense to us or say things that seem irrational. Yet our ability to perceive the everyday world of social action as a world of meanings, purposes, intentions and reasons is an extraordinary one.

At the heart of the ability to read meaning in perceived action is *parsing*. A characteristic of skilled action is that, in physical terms, no structure is overt. The

R.W. Byrne (✉)

School of Psychology, University of St Andrews, St Andrews, Fife KY16 9JP, Scotland
e-mail: rwb@st-andrews.ac.uk

sequence of components is *linear*—although with action that involves both hands or extends to mouth or feet, there may be several partly-linked streams running in parallel, each linear in sequence. But whether driving a car, uttering a sentence, or baking a cake, all that is physically present to be perceived is *smooth, fluid movement*. The absence of “real gaps” between many of the separate words in a spoken sentence is part of every entry-level linguistics course; and just the same is true of manual actions. Once a skilled sequence of actions has been assembled, practising will result in smoother and smoother performance, to the point when underlying structure is not signalled by any detectable interruptions in the sequence. That is the first part of the parsing problem: seeing a linear sequence of fluid behaviour, but perceiving it as segmented into discrete units that correspond to real entities for the actor who is observed.

The second parsing problem concerns the fact that organized, complex behaviour is *hierarchical* in structure. This means that elements lying together in sequence may be closely related logically, because they form part of a module or subroutine or phrase (depending on what sort of behaviour is under discussion); or much less closely related, only lying together by virtue of the organization of some higher order unit of organization. To understand action, and thereby detect the meaning in it, it is crucial to parse its hierarchical structure accurately. The output of the parsing process must go beyond a sequence of discrete units, to get at the underlying *relationships* that we conventionally represent in terms of a bracketed string, a tree-diagram or a phrase-structure grammar. Without that, there would be no systematic way to connect observed behaviour to the purposes that underlie it in the mind of the actor—and thus, to go on to understand the actor’s *intentions* and the *cause-and-effect* of how that particular behaviour is efficient for achieving their purposes.

It is the thesis of this chapter that parsing has its evolutionary origins in an unexpected place. Rather than deriving from a selective pressure for more sophisticated vocal communication, a function in which we see the full flowering of parsing ability in modern humans, I argue that parsing was originally part of a *feeding adaptation*; and that these abilities, evolved for efficient feeding, were themselves based on earlier evolution of abilities in *social behaviour reading*.

After briefly considering primate vocal communication, I will first sketch the evidence that a segmentation system, one that can parse a smooth behavioural performance into separate but meaningful units of action, is present in monkeys—and probably in many other species even more distant from us on an evolutionary time-scale. The main biological function of action segmentation in those species is most likely the estimation of current behavioural dispositions in conspecifics and the prediction of their likely actions in the immediate future. Among the primates, it seems, only in great apes did rather special abilities of *hierarchical* parsing and anticipatory planning develop, and I will suggest that these special capacities may be parasitic on that earlier segmentation system—but are not dependent on any prior ability to understand intentions or causality. In non-human great apes hierarchical parsing seems only to be only found within the *manual skill domain*, where it functions in the wild by allowing more efficient feeding; and there are plausible ecological reasons why enhanced feeding abilities should have evolved specifically in the

great apes. Under the artificial conditions of human rearing, hierarchical parsing and anticipatory planning give rise to a wide range of richly complex behaviours, and can be deliberately co-opted into human-derived communication systems such as American Sign Language.

Given such abilities in living apes, in the manual-spatial domain, it is only a small step to speculate that in one of our own early ancestors these hierarchical parsing skills became available also in the vocal-auditory domain. Linguistic syntax is thereby seen as evolutionarily derived from hierarchical behaviour parsing. Further implications may be drawn out. As emphasized, behaviour parsing is not dependent on prior causal-intentional understanding; however, it could have been a crucial step on the way to achieving this level of mental representation—an essential precursor to human cognition, and still a necessary part of the process of representing phenomena as causal-intentional structures. Moreover, the fact that so much can in principle be achieved *without* involving that level of mental representation—parsing of behavioural structure, social learning of complex skills by program-level imitation, and so on—opens the door to a heretical thought. Could it be that the prevalence of causal-intentional interpretation of our social world is illusory, a consequence of retrospective contemplation? Certainly, when we choose to ponder causation and attribution, or when we are asked to justify our actions by others, as adult humans we are well able to construct causal-intentional theories that make sense. But perhaps the cut-and-thrust of everyday social action and interaction does not need this mentalizing, or would indeed be slowed or disrupted by it (Apperly & Butterfill, 2009; Bargh & Chartrand, 1999), and we should look elsewhere for the evolutionary functions of theory of mind and causal reasoning.

8.1 Primate Vocal Communication—Primitive Speech?

Extensive study for many years has focused on primate vocalizations, driven partly by theoretical interest in language origins and partly by the availability of sound-manipulation technology. We now know that the potential for flexibility in the production of calls by primates is very limited.

No primate can copy another's sounds, in the way that many birds and some cetaceans can do (Janik & Slater, 1997). Even vocal dialects are nearly unknown in primates, except in cases where human influence may have unintentionally conditioned a local variation (Green, 1975; Mitani, Hasegawa, Gros-Louis, Marler, & Byrne, 1992). “Nearly”, because there is recent evidence that zoo communities of chimpanzees develop characteristic group dialects (Auser & Wrangham, 1987), and adjacent communities in the wild have been found to differ more in their vocalizations than do more distant ones—just as a dialect in human communities can serve to identify group membership and label an out-group (Crockford, Herbinger, Vigilant, & Boesch, 2004). Even in these cases, the modifications are small ones, to calls which are biologically fixed in form. Young primates of many species have often been reared out of any auditory contact with conspecifics: nevertheless, they

all develop a normal repertoire of vocalizations. Learning does play a role in the normal development of calling, but this is *contextual learning* not production learning (Janik & Slater, 1997): primates learn the appropriate circumstances in which to call, rather than learning the calls themselves. The famous case of predator-specific alarm calls in vervet monkeys shows this process in action (Seyfarth & Cheney, 1986). The referential specificity of these calls is to a limited extent innate, but whereas a young vervet will initially make an “eagle alarm” to a wide range of flying things (even a large, falling leaf on one occasion), as it matures calling is restricted to large broad-winged birds, then specifically to raptorial species, and finally the call is given almost exclusively to the martial eagle *Polemaetus bellicosus*, a vervet’s main aerial predator.

Most non-human primates have a vocal repertoire of more-or-less discrete calls, but also show some graded variation, most extensively in the chimpanzee and gorilla (Marler & Tenaza, 1977). Animals have been found to perceive human speech categorically (Kuhl, 1982), and primate calls which sound like a smoothly varying continuum to the human ear have been shown to be composed of several circumstance-specific and function-specific calls (Gouzoules, Gouzoules, & Marler, 1984). However, nothing remotely like the multiple levels of patterning and syntactic structuring found in human speech has been detected in any primate vocal system. The closest to hierarchical organization is the recent discovery that one call can modify another and so qualify its degree of definiteness, as if adding “maybe” to its meaning (Zuberbuhler, 2002). This is a far cry from the generative, productive nature of everyday human speech, and theories that try to make direct connection between primate vocal communication and language have a large gap to fill—with pure speculation. For these reasons, I now turn instead to the manual skills of human and non-human primates.

8.2 Segmentation of the Action Stream

When we approach a range of problems, from car maintenance to public speaking, we do so with a pre-existing repertoire of motor actions ready to deploy. Some of these “elements” of action are no doubt innate; and many others are constructed by trial-and-error exploration of previous similar situations; and a significant part of our repertoire of actions is learned by noticing other people’s behaviour (and listening to their speech). By all three routes of acquisition, we meet the many novel problems in adulthood prepared with a rich vocabulary of elements of action which we can permute and organize into creative solutions, as well as deploy effortlessly in response to more familiar demands.

The stream of action that we observe, however, does not come with ready-made gaps that correspond to logically distinct elements. This has been classically noted to apply to speech, where a sound-gap is more likely to be part of a plosive consonant than to signal a new word, but in fact the point applies to all skilled behaviour. Thus, with motor action, the physical stimulus that confronts us is smooth and fluid, not

segmented. How are we nevertheless able to pick out functional elements in the smooth and apparently unbroken flow of action?

To be used as building blocks in effective motor planning, elements of action discerned in another's behaviour must meet one simple principle: each element should *already* be within the repertoire of the observer (Byrne, 2003). In contrast, the "size" of an element is irrelevant: I propose that people are able to "see" (pick out) within a stream of action any element which is already present as a pattern in their personal repertoire. For different observers, or at different times in the life of a single observer, one particular movement of a single finger or an elaborate sequence of bimanual movements might both properly be seen as single elements. When we watch a relatively unfamiliar process being performed, the *level* at which we notice elements will be low, perhaps that of finger movements; whereas when we watch a slight variant of an already familiar activity, the basic elements that we notice might themselves be high-level, complex processes. Most commonly perhaps, the level at which observed behaviour matches parts of our existing repertoire would be neither of these, but rather consist of simple and highly-practised movements that produce visible effects on environmental objects: that is, simple, goal-directed movements. Such elements may be particularly easy to delimit because they are marked by a characteristic pattern of acceleration and deceleration, just like the cadence of syllables in a sentence. Consistent with this idea, people are able to pick out the basic elements of action, even when the stimulus is experimentally reduced to fluorescent spots on the joints (Baldwin, Andersson, Saffran, & Meyer, 2008; Loucks & Baldwin, 2009). Is it plausible that this means of segmentation is a primitive part of the human cognitive system? A digression into recent neuropsychological studies of monkeys suggests that it is.

Non-human primates have been shown able to pick out, in the behaviour of others they observe, actions that are already in their own repertoire. A system of single neurons has been identified in the premotor cortex of rhesus monkeys *Macaca mulatta* (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Fogassi, & Gallese, 1996, 2002), each of which responds to a simple manual action, and responds equally whether the monkey makes the action or sees another do it. The cardinal properties of these *mirror neurons* are (1) they detect goal-directed movements that are in the observing monkey's own repertoire, and (2) they generalize over whether the movement is performed by the monkey itself or by another agent. It is unlikely that mirror neurons have any role in imitation for monkeys, simply because monkeys have repeatedly failed to show evidence of imitative capacity (Visalberghi & Fragaszy, 1990). Rather, Rizzolatti and his collaborators relate the evolutionary origin of mirror neurons to monkey social sophistication: i.e. they suggest that the system functions in revealing the demeanour and likely next actions of conspecifics, by reference to those actions the observing monkey might itself have done (Rizzolatti et al., 2002).

These units have sometimes been described as "monkey see, monkey do" cells, and in a very restricted sense this is accurate. Much of what is described as imitation in experimental studies of non-human primates involves provoking a subject to repeat an action that is in its repertoire, upon seeing another perform the same

action (e.g. Whiten, Custance, Gomez, Teixidor, & Bard, 1996; Custance, Whiten, & Fredman, 1999; Bugnyar & Huber, 1997; see Byrne, 2002a for discussion). However, in these studies nothing new is being learned: this sort of imitation has been argued to be better described as *response facilitation* (Byrne, 2002a, 2002b; Rizzolatti et al., 2002). In response facilitation, as opposed to any more general sense of imitation, a pre-existing response is facilitated (i.e. made more available) by seeing it done, and this causes a higher probability of the response occurring subsequently (Byrne, 1994; Byrne & Russon, 1998). Response facilitation is closely related to stimulus enhancement (Galef, 1988; Spence, 1937), and they may indeed be two manifestations of the same phenomenon: *priming* of neural correlates (Byrne, 1994, 1998b, 2005a). On this view, priming neural correlates of some aspect of the social situation or the environment results in stimulus enhancement; whereas priming neural correlates of an action pattern within the current repertoire results in response facilitation. The mirror neuron system provides a possible neural instantiation for imitation in this restricted sense of response facilitation, but not for imitative learning of new skills. (This does not mean the mirror system is innate and fixed; indeed, at least for the human equivalent, there is evidence that it can be affected by learning: Catmur, Walsh, & Heyes, 2007.)

Despite these reservations, a segmentation system, based on elements of action that the observer can already perform, would be a very useful starting point for more elaborate forms of imitation—and that is what I have proposed underlies great ape imitation (Byrne, 2003). By responding to precisely those movement patterns that correspond to potential actions in the viewer's repertoire, segmentation by response facilitation, operating by means of mirror neurons, has the power in principle to convert automatically a continuous flow of observed movements into a *string* of recognized, familiar actions. If seeing a string of familiar actions also allows construction of links between them, then “action-level” imitation occurs (Byrne, 2002a, Byrne & Russon, 1998). In action-level imitation, a linear sequence of actions is copied without recognition of any higher-order organization that may be present: the organization is “flat”. Chimpanzees have been reported to copy the order of actions, even though the sequence was entirely arbitrary and unrelated to success (Whiten, 1998), and a detailed learning model has been developed to describe action-level imitation in animals (Heyes & Ray, 2000).

If it were beneficial to copy arbitrary, random actions or behaviour that is genuinely linear in structure (e.g. the “fixed action patterns” described by early ethologists), action-level imitation might be useful. However, most human action, and arguably also much of the behaviour of non-human great apes, is planned: with a hierarchical, not linear organization. The question is, can this planning be “seen” by an ape, in the behaviour of another? More generally, can a bottom-up, mechanistic analysis go beyond action-level imitation to explain how behavioural organization can also be parsed and thereby copied, i.e. *program-level imitation* (Byrne & Russon, 1998)? If so, then the evolution of behavioural parsing has implications far beyond imitation itself.

8.3 Parsing Hierarchical Structures of Behaviour

It is no coincidence that a theory of how program-level imitation might be achieved should have been developed to explain manual behaviour, specifically in great apes (Byrne, 2003). Most animals simply do not learn sufficiently complex patterns of behaviour for imitative learning to be detectable in observational data from them, nor would they have much need for the ability to learn by imitation (Byrne, 2002a). The primitive 5-fingered primate hand (Napier, 1961) is highly effective as a manipulator and in many species shows some opposability. In great apes, however, the hand shows a considerably augmented range of aptitudes compared even to those of monkeys. For example, in the mountain gorilla (Byrne, Corp, & Byrne, 2001b), everyday food preparation typically involves using the two hands in different but complementary roles (i.e. *manual role differentiation*: Elliott & Connolly, 1984). The resulting “asymmetric bimanual co-ordination” is augmented by the gorilla’s ability to control individual digits of the hand independently (i.e. *digit role differentiation*: Byrne et al., 2001b). This allows items to be held in part of the hand while other digits can carry out other activities; for instance, part-processed food can be accumulated in the hand, while part of the food-processing routine is iteratively repeated to build up a larger handful of food. This remarkable dexterity allows mountain gorillas to deal with plants that are physically defended by an array of spines, stings and hard casings (Byrne, 2001). In the process, they display a huge repertoire of functionally distinct elements of action (i.e. single actions that produce clear changes to the plant substrate; for instance, thistle-processing alone requires 72 such elements). Although attention has been drawn away from the chimpanzee’s general manual skills by the anthropological emphasis on tool use, when chimpanzee plant-processing has been studied qualitatively, similar abilities are found to those of gorillas (Corp & Byrne, 2002a, 2002b). (Note that the sophisticated tool-making of early hominins also relied on manual dexterity: Byrne, 2004, 2005b). With animals of such dexterity, manual behaviour is sufficiently rich for complex organizations of learnt behaviour to be detectable by researchers. Moreover, the manual tasks confronted by great apes are challenging, so it would certainly pay the apes to be able to learn new skills by imitation of others.

The evidence that great apes do indeed learn skills by imitation comes from observational data rather than experiment, since no useful experimental test of program-level imitation in animals has yet been devised. Although the evidence is therefore oblique, cumulatively it is fairly impressive (Byrne, 2002a, 2005a). First, there is the very fact that young great apes learn complex, hierarchically structured routines of manual behaviour (some of them essential to survival in adulthood) in just a few years before their weaning, in contrast to monkeys where there is no evidence of anything comparable. Evidence of complexity is strongest for the mountain gorilla, where 5-stage sequential processes have been described (Byrne, 1999c, Byrne & Byrne, 1993; Byrne, Corp, & Byrne, 2001a), but also clear in chimpanzees, both in tool-using tasks (Boesch & Boesch, 1990; Goodall, 1986; Matsuzawa, 2001; Matsuzawa & Yamakoshi, 1996) and in dealing with complicated

plant foods (Corp & Byrne, 2002a; Stokes & Byrne, 2001). The fact that orangutans sometimes also use tools to deal with complex plant defences (Fox, Sitompul, & Van Schaik, 1999) suggests that they have similar abilities, and this is confirmed by studies of young orangutans' efforts to deal with the vicious spines of certain palm trees (Russon, 1998). Far more studies have been carried out on the foraging behaviour of monkeys than that of apes; yet no comparable evidence has come to light. Second, in a detailed analysis of variation in the skills of adult mountain gorillas, it was striking that minor details (grip type, exact fingers employed, hand preference, extent of movement) varied idiosyncratically between individuals, even between mother and offspring, whereas the overall "program-level" organization of each technique was remarkably standardized in the local population (Byrne & Byrne, 1993). If idiosyncrasy is characteristic of trial and error learning, such standardization of techniques needs explaining. There are two possibilities: either the affordances of the gorilla's hands, combined with the physical form of the plant defences, define a clear gradient of optimization and thus with practice every gorilla will inevitably acquire the same method; or, observational learning is involved, and some aspects of the skills are passed on culturally. The third line of evidence is specifically relevant to this issue, as it involves the study of animals disabled by crippling snare wounds. Snares are not set to catch gorillas, but young individuals may suffer injury because of their explorative behaviour (Stokes, Quiatt, & Reynolds, 1999). If the standardized pattern of an adult technique is a product of affordances, then in an animal with severely maimed hands a quite different technique should result from the same trial and error experience. Yet in both chimpanzees and gorillas, disabled individuals acquire the same organization of behaviour as the able-bodied, and instead work around their difficulties by modifying the low-level details of implementation (Byrne & Stokes, 2002; Stokes & Byrne, 2001). This favours the hypothesis that the standard technique is a culturally transmitted pattern. Finally, one anecdotal observation supports the case that great apes can only learn certain aspects of their complex feeding skills by observation. When processing stinging nettles, one single adult in the study population differed in technique: the female Picasso did not fold bundles of leaves, so was presumably often stung on her lips (Byrne, 1999a). Picasso had transferred into the study area from lower altitude, where nettles do not grow. Because adult gorillas feed alone and out of sight of others in dense herbage, mountain gorillas' only opportunity for observational learning of plant processing comes in infancy. It seems most likely that a lack of opportunity to observe accounts for Picasso's incomplete technique, and intriguingly her juvenile was the only other gorilla in the study population to lack that particular element of the skill.

8.4 Imitation Without Intentionality

In the face of this evidence, I therefore developed a theory of how great apes could learn the program-level structure of behaviour by imitation, aiming to avoid any assumption that the animals had prior understanding of purpose or intention (Byrne, 1999b, 2003). This "behaviour parsing" model is based instead on the *statistical*

regularities present within the variability of multiple performances of the same skilled sequence of action.

Every execution of a motor act, however familiar and well-practised it is, will differ slightly from others. Nevertheless, this variation is constrained—because if certain characteristics are missing or stray too far from their canonical form the act will fail to achieve its purpose. Watching a single performance will not betray these underlying constraints, but the statistical regularities of a repeated, goal-directed action can serve to reveal the organizational structure that lies behind it. Unweaned great apes spend most of each day within a few feet of their mothers, and (since their main nutrition still comes from milk) they have almost full-time leisure to watch any nearby activities, as well as learn about the structure of the local environment by their own exploration. For instance, by the time a young gorilla first begins to handle a plant like a nettle, at the rather late age of about two years because the stinging hairs discourage earlier attempts, it will have watched many hundreds of nettle plants being expertly processed by its mother.

Consider how a young gorilla might learn from statistical regularities of observed behaviour how to process stinging nettles (Fig. 8.1).

Its mother's behaviour will be perceived as a string of discrete elements, where each of these actions is a familiar one that it can already perform. At this time, the young ape's repertoire of familiar elements of action derives from: (i) its innate manual capacities; (ii) from many hours of playing with environmental objects, such as plants and discarded debris of the mother's feeding; and (iii) from its own experience of feeding on other plants, perhaps ones simpler to process than nettles. Suppose that it also has some way of focusing on those particular sequences of its mother's action that are relevant to eating nettles; perhaps it has explored nettle plants and found that they are painful, yet puzzlingly the mother seems to enjoy interacting with them, making her nettle-interactions intrinsically interesting. (Some such mechanism to focus learning on relevant action sequences seems to be essential for any "bottom up" model of motor learning.) Because motor behaviour is intrinsically variable, and plants also vary from individual to individual, the string of elements that the young gorilla sees when watching its mother eat nettles will differ each time. However, her starting point will always be a growing, intact nettle stem, and—because she is expert at this task—her final stage will always be the same, popping a neatly folded package of nettle leaves into the mouth. In between these points, variation will be particularly associated with non-critical parts of the performance, and certain aspects must necessarily be the same—or else, the result simply will not be success. With repeated watching, and a mind that tends automatically to extract regularities in behaviour that varies over time, a pattern will gradually begin to become apparent. The mother always makes a sweeping movement of one hand, held around a nettle stem which is sometimes held in the other hand even though the plant is still attached to the ground, and this leaves a leafless stem protruding from the ground; she always makes a twisting movement of the hands against each other, and immediately drops a number of leaf-petioles (which she does not eat) onto the ground; she always uses one hand to fold a bundle of leaf-blades protruding from

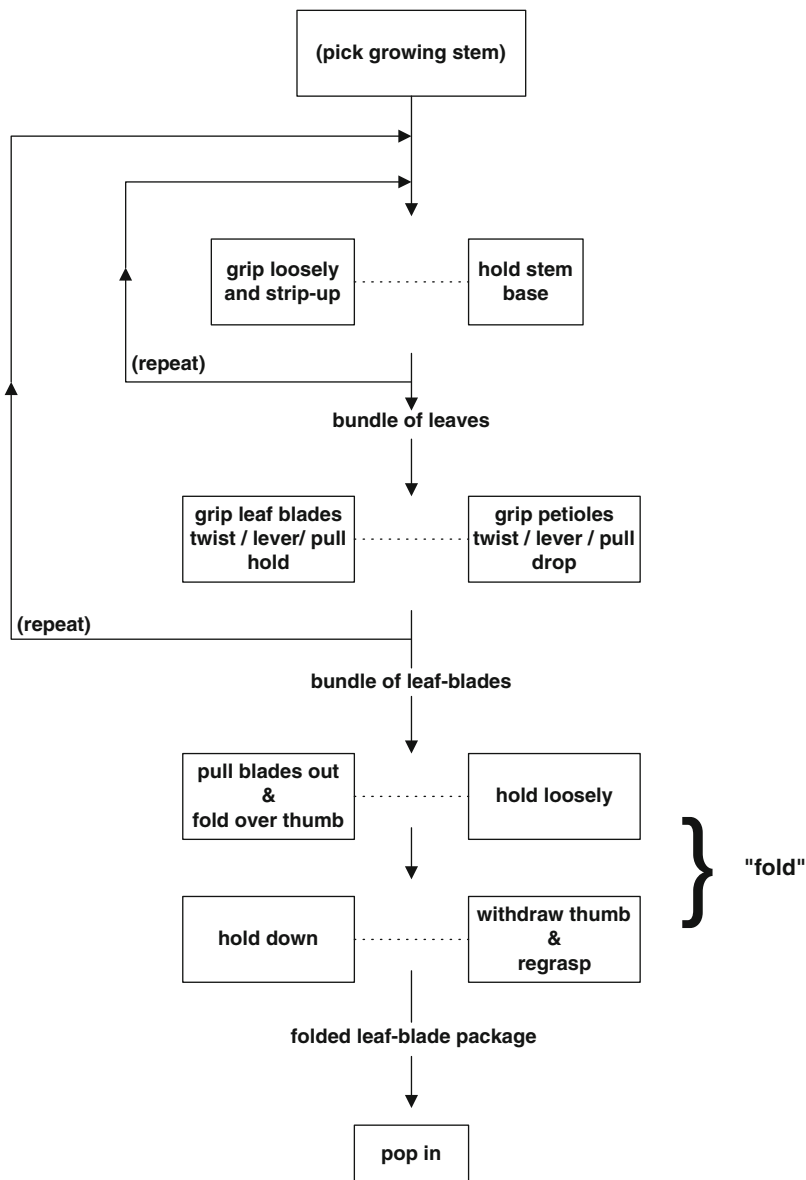


Fig. 8.1 Flow-chart for a typical adult gorilla processing nettle *Laportea alatisipes* leaves. The action starts at the top, with selection of a growing nettle to eat, and works downwards. Processes are shown in rectangles; those which are optional, depending on the state of the plant itself, are shown in brackets. As with conventional flow-charts, diamonds represent choice points, with the alternative options shown by the directed links leading from each diamond. Unlike the single linear process of most flow-charts, the diagram represents the actions of both *left* and *right hand*: actions which are significantly lateralized to the *left hand* are shown on the *left* of the figure, and vice versa for the *right hand*. Some of these actions are nevertheless co-ordinated together, though the two actions are different: these cases of asymmetric bimanual co-ordination are shown with *dotted lines* connecting the separate processes

the other hand, and holds down this folded bundle with her thumb. Moreover, these stages always occur in exactly the same order each time.

Statistical regularities, in behaviour that is repeatedly observed, thereby mark out the minimal set of *essential actions* from the many others that occur during nettle eating but which are not crucial to success, and reveal the *correct order* in which they must be arranged. (The ability of human babies as young as eight months to detect statistical regularities in spoken strings of nonsense words shows that just such sensitivity to repeated orderings is active early in human development: Saffran, Aslin, & Newport, 1996.). The usefulness of detecting regularities applies not only to the linear sequence of movements of each hand, but also the hands' operation together: stages that crucially depend on the hands' close temporal and spatial co-ordination while doing different jobs will recur in every string, while other coincidental conjunctions will not.

Other statistical regularities derive from modular organization and hierarchical organization. Whenever the operation of removing debris is performed (by opening the hand that holds nettle leaf-blades, and delicately picking out debris with the other hand), it occurs at the same place in the sequence. Also, on some occasions but not others, a section of the program sequence may be repeated twice or several times. For instance, the process of <pull a nettle plant into range, strip leaves from its stem in a bimanually co-ordinated movement, then detach and drop the leaf-petioles> may be repeated several times before the mother continues to remove debris and fold the leaf-blades before eating. Subsections of the string of actions that are marked out in this way may be single elements, or as in this example a string of several elements. Both omission and repetition signal that some parts of the string are more tightly bound together than others, i.e. that they function as *modules*. Optional stages, like cleaning debris, occur between but not within modules. Moreover, repetition of a sub-string gives evidence of a module used hierarchically as a *subroutine*, for example, iteration to accumulate a larger handful.

Further clues to modular structure are likely to be given by the distribution of pauses (occurring between but not within modules), and the possibility of smooth recovery from interruptions that occur between modules. Gorillas often pause for several seconds during the processing of a handful of plant material, in order to monitor the movements and actions of other individuals. Finally, a different module entirely may be substituted for part of the usual sequence (e.g. if one hand is required for postural support, then a normally bimanual process may need to be performed unimanually), and if this module is recognized as an already-familiar sequence its substitution again reveals structure; eventually, it may be that a taxonomy of substitutable methods is built up.

All these statistical regularities are precisely what enabled us, the researchers, to discover the hierarchical nature of nettle processing by adult gorillas (Byrne & Byrne, 1993). The behaviour parsing model proposes that the same information can be extracted and used *by the apes themselves*, and that this ability is what enables a young ape to perceive and copy the sequential, bimanually co-ordinated, hierarchical organization of complex skills from repeated watching of another.

Behaviour parsing enables the underlying hierarchical organization of planned behaviour to be picked out—under certain circumstances. The first caveat, from what we know of living apes in the wild, is that it is entirely possible that non-human apes' capacity to parse behaviour is limited to the visible domain of manual and bodily actions, and thus not available in the auditory domain. The bonobo Kanzi's apparent ability to parse human speech, when he responds correctly to words whose referent depends on the syntactical organization of a relative clause within a sentence (Savage-Rumbaugh et al., 1993), may cause this qualification to be relaxed, at least for extensively human-reared apes. For the moment, however, I will assume that living apes under natural conditions, and our own earliest ancestors, had no such ability. The great ape forte is evidently the manual domain, as convincingly demonstrated in the hundreds of ASL signs acquired by participants in "ape language" experiments (see chapters in Gardner, Gardner, & Van Cantfort, 1989). In contrast, modern humans are routinely able to parse vocal material.

The second limitation, from the way the model works, is that "multiple independent looks" are necessary. A single view of skilled behaviour that is unfamiliar in its organization will not result in a useful parsing, so seeing *multiple* samples of efficient behaviour is required. The samples must be *independent*, so that there is information about the variance within the strings of perceived elements; that is because only by having sensitivity to the relative variability of elements can behaviour parsing locate the key (unvarying) elements. Thus, viewing a film-clip of the same segment of skilled behaviour would not serve to allow unfamiliar behaviour to be parsed. Note that, although we may well substantially overrate our everyday abilities (Bargh & Chartrand, 1999), modern humans do not seem to be subject to this limitation. Gergely, Bekkering, and Kiraly (2002) show that babies over sixteen months old are able to pick out for imitation the key elements of behaviour demonstrated only once, according to simple rationality criteria; behaviour parsing alone could not explain these data. Before the critical age, I predict that babies are still able to show program-level imitation, but will not at that point be able to select out specifically rational features of the process to copy. In circumstances not requiring acquisition of new behavioural organization, there is also some evidence for similar selectivity in imitation without multiple views of the behaviour in both chimpanzees and domestic dogs (Buttleman, Carpenter, Call, & Tomasello, 2007; Horner & Whiten, 2005; Range, Viranyi, & Huber, 2007). Thus, the behaviour parsing model can only be part of the eventual answer of how human imitative abilities evolved.

8.5 Why Great Apes?

To be precise, why should it have been only this one taxon of primate that developed the rather special ability to parse a segmented stream of action into a hierarchically-organized structure—and thereby acquire novel, complex skills by imitative learning? At present, the social brain or Machiavellian intelligence

hypothesis is widely accepted as the most plausible explanation for the origin of primate intelligence (Brothers, 1990; Byrne & Whiten, 1988; Dunbar, 1998; Humphrey, 1976; Jolly, 1966; Whiten & Byrne, 1997). However, when it comes to accounting for cognitive differences *between* monkeys and apes, it will not do. According to the social brain hypothesis, the root cause of intellectual advance is social complexity. Because the ancestors of modern haplorhine primates (monkeys and apes) needed to live in increasingly large social groups, yet individuals of each species were thereby put in direct competition for resources with other group members, a selection pressure resulted that favoured increased social intelligence and a concomitant enlargement in neocortex volume (Byrne, 1996). Thus, today, we find that primates living in larger groups have larger brains (Barton & Dunbar, 1997; Dunbar, 1992), and are more likely to employ subtle means of social manipulation such as deception (Byrne & Corp, 2004). While this fits nicely with the differences among living species of varied brain sizes, and gives a good account of the evolutionary origins of the large-brained haplorhines, it does not distinguish between monkeys and apes. There is no systematic difference in the causal variable: the great apes simply do not live in larger social groups than do many monkey species, which have much smaller brains and show little sign of the sophisticated cognition of apes.

This means that serious attention must be paid to alternative, ecological selection pressures that might have promoted intelligence, at least for this special case (Byrne, 1997): for instance, is there an ecological challenge that affects great apes more than monkeys? Because of the anatomical differences between monkeys and apes, the answer is yes. Great apes are systematically larger than monkeys, and since they are adapted to brachiation (hanging below branches on long, powerful arms) costs of long-distance travel are much greater for them than for monkeys. However, apes are all specialists in easy-to-digest plant material (fruit or soft leaves) which is ephemerally available and patchily distributed, so apes they must regularly travel to find their food. Almost everywhere they live, great apes share the forest with Old World monkeys—which are not only smaller and more efficient in long-range travel, but happen to have gut adaptations enabling them to eat fruit when slightly less ripe, or leaves when slightly tougher, than can apes. Monkeys, in short, are in direct niche competition with great apes and possess all the aces: how have living apes survived at all? The explanation becomes clear when the details of their diet are examined: chimpanzees make tools to extract social insects from their nests, and to break open hard nuts; gorillas, and to a lesser extent chimpanzees, use elaborate, multi-stage routines to deal with plant defences; orangutans use complex, indirect routes to reach defended arboreal food, and sometimes make tools to gain access to bees' nests or defended plant food. In each case, "clever" methods of food extraction are used to gain access to foods which monkeys would be unable to reach. Thus, it becomes plausible that the Miocene ancestors of the living great apes (whom we share) may have adapted cognitively, in ways that would enable a broader range of food types to be exploited: and I propose that learning new skills by behaviour parsing was just this adaptation.

8.6 Parsing to “See” Intentions: The Origin of Mime and Gestural Language?

If this behaviour parsing model is correct, human language and speech evolved in a species that was *already* able to parse hierarchically organized behaviour—which might be no coincidence. Moreover, this ability to “see below the surface” of behaviour, and detect the logical organization that produced it, has implications for other cognitive activities. Indeed, the ability to learn new skills by imitation may be seen as just part of a fundamental process of interpreting or understanding complex behaviour.

It was important in the development of the behaviour parsing model that processing should start from observed behaviour and require no prior understanding of the physical cause-and-effect of the actions upon objects in the world, nor the intentions or other mental states of the demonstrator. However, we know from common experience that these more abstract representations form regular parts of how adult humans understand and discuss the world: so their evolutionary origin must be explained. Behaviour parsing might be a *necessary* step on the road to seeing the world in an intentional-causal way.

Consider causation. Since a perceptual parsing of complex action will (in many cases) be applied to actions-upon-objects in the world, changes in the physical world will become linked to the sequence of action—statistically. Of course, there is more to cause than correlation, but it can be questioned whether that matters for everyday purposes, or for evolution. Reliable correlation might be described as a “Pretty Good Cause”, and only physicists dealing with the fundamentals of matter may need to go much beyond it. The fact is that most things are seen as likely to happen to the extent that they, or things very like them, have happened before under the same circumstances. The sun will rise tomorrow morning because it has been doing so for a long time at rather regular and statistically predictable intervals; not flawless logic, but good enough. Any parent who has tried to answer a series of “Why?” questions from a young child will know how soon one gets out of one’s depth with causation: ok, so day and night are caused by the Earth going round the Sun, but why does it do that? In fact, probing deeper into the physics of most everyday situations helps little with everyday living, and does not provide a very satisfying advance on *cause-as-correlation*. In contrast, behaviour parsing picks out the correlational structure of a changing environment quite well.

How could behaviour parsing help us with intentionality? The perceived organization of behaviour that results from the parsing process will inevitably be set in a real-world context of achievement of valuable ends, just because the individuals observed engaged in skilful action will only be doing so for biologically sufficient reasons. Often, demonstrators will be close associates or relatives of the observers, confronting much the same problems as them. Thus, associating a particular organizational structure with the typical *result* of its performance is in many cases a relatively trivial task: the *point* of achieving that particular result is something the observer probably already understands. Intended purpose is indicated by the usual

result of successful performance. (“Unsuccessful” is of course also identified statistically, here, on the basis of visible behaviour. It corresponds to those occasions when the action needs to be re-done, rather than moving on to another action.) This means that, in principle, behaviour parsing makes it possible to compute the prior intention of the other individual: by recognizing a behaviour pattern that would, if the observing self performed it, achieve a comprehensible goal for the self. Any animal capable of program-level imitation should therefore also be able to detect at least some intentions of others from their behaviour, in cases where they have been able to gain the necessary prior experience of that behaviour. And indeed, the living great apes do show some aspects of theory of mind (Byrne, 1995, 1998a, 2000; Cartmill & Byrne, 2007; Tomasello, Call, & Hare, 2003), although it seems likely that these fall short of the full mentalizing abilities of five year old children (Astington, Harris, & Olson, 1988; Perner & Wimmer, 1985; Wellman, 1990). As in the case of causation, the intentions extracted by behaviour parsing are intentions in a weak sense of the term: rather than an imagined mental state, intentions of these kinds need be no more than proper results of the normal behaviour sequence. But similarly, this sense of intention may be good enough for most everyday purposes: animals sensitive to *intentions-as-results* will not be able to conceive of false belief and deliberate trickery, but they will be able to pick out the purposes of many everyday social actions.

Animals with behaviour parsing abilities, as indexed by their ability to imitate at program-level, might still be rather limited in understanding—with causation reduced to correlation, and intentions reduced to expected results. However, combined with the delicate and sophisticated manual control of action that we find in all the living apes, even this limited kind of understanding should be sufficient for communication by means of *gesture*. Natural gestural communication in non-human apes is a rather neglected topic, but current evidence shows that in captivity both chimpanzees and gorillas develop gestures not seen in the wild, and use them intentionally in dyadic communication (Genty & Byrne, 2010; Genty, Breuer, Hobaiter, & Byrne, 2009; Hobaiter & Byrne, 2010, 2011; Pika, Liebal, & Tomasello, 2003; Tanner, 1998; Tanner & Byrne, 1996, 1999; Tomasello, George, Kruger, Farrar, & Evans, 1985; Tomasello, Gust, & Frost, 1989). Moreover, the ability of living great apes to extend their gestural repertoires when helped by humans has been amply demonstrated in the various “ape sign language” projects: whatever is believed of their linguistic sophistication, there is no doubt that those chimpanzees, gorillas and orangutans have learned many new manual gestures.

8.7 Tailpiece: A Heretical Thought

Those who conduct behavioural experiments or analyse observational data from the field, in order to discover whether any animal has the ability to represent the mental states of others, become acutely aware that their task is a difficult one

because simpler mechanisms can generate richly complex behaviour. In particular, this chapter has argued that an understanding of planned behaviour, in terms of hierarchically organized structure that can be copied, with causality approximated by correlation and purpose by normal results, can result from a mechanistic process of behavioural analysis that need not involve any “mentalizing” about the actual mental states of the observed party. Thus, great apes show program-level imitation, but might still not possess theory of mind and causal understanding. But what about humans?

Of course, humans can and do represent causes and intentions: we explain (away) our actions, on grounds of our beliefs, false or otherwise; we teach our children by explaining that one thing causes another or that some people have different beliefs to ourselves, and so on. But do these retrospective, verbal accounts actually correspond to causal mental states that generate our behaviour when we are not explaining anything? We are always reluctant to accept how much of our behaviour is an automatic and fast product of mental processes of which we are unaware (Bargh & Chartrand, 1999), but I think this should be seriously considered for the case of theory of mind.

There are two possibilities. On the one hand, it may be that calculations about others’ mental states *are* causal, and that the normal process of automatization with practice simply renders them faster and more efficient, to the point when they can only be made conscious by “off-line” deliberation. But the heretical alternative is that rather different, mechanistic but unconscious processes—analogueous to those that allow us to parse behaviour—actually cause most of our everyday social behaviour and interactions with the world of objects, and mentalizing is a secondary process (and see Apperly & Butterfill, 2009 for a related discussion). On this view, mentalizing has different functions: these include teaching, when we explain processes or people to a child, and prevaricating, when we retrospectively construe our behaviour in a way very different to what we know to be accurate. Any such process of verbal (mis)construal is certainly a function of language ability, and so must be recent in human evolution; but it may be that the behavioural capacities that we attribute to “theory of mind” were all present at an earlier stage in human evolution, and are perhaps even shared with non-human great apes, though they cannot explain and discuss their actions as we can.

References

- Apperly, I., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*, 953–970.
- Astington, J. W., Harris, P., & Olson, D. R. (1988). *Developing theories of mind*. Cambridge: University Press.
- Auser, M. D., & Wrangham, R. W. (1987). Manipulation of food calls in captive chimpanzees. *Folia Primatologia*, *48*, 207–210.
- Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition*, *106*, 1382–1407.
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, *54*, 462–479.

- Barton, R. A., & Dunbar, R. I. M. (1997). Evolution of the social brain. In A. Whiten & R. W. Byrne (Eds.), *Machiavellian intelligence II: Extensions and evaluations* (pp. 240–263). Cambridge: Cambridge University Press.
- Boesch, C., & Boesch, H. (1990). Tool use and tool making in wild chimpanzees. *Folia Primatologica*, *54*, 86–99.
- Brothers, L. (1990). The social brain: A project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience*, *1*, 27–51.
- Bugnyar, T., & Huber, L. (1997). Push or pull: An experimental study of imitation in marmosets. *Animal Behaviour*, *54*, 817–831.
- Buttleman, D., Carpenter, M., Call, J., & Tomasello, M. (2007). Enculturated chimpanzees imitate rationally. *Developmental Science*, *10*, F31–F38.
- Byrne, R. W. (1994). The evolution of intelligence. In P. J. B. Slater & T. R. Halliday (Eds.), *Behaviour and evolution*. Cambridge: Cambridge University Press.
- Byrne, R. W. (1995). *The thinking ape: Evolutionary origins of intelligence*. Oxford: Oxford University Press.
- Byrne, R. W. (1996). Machiavellian intelligence. *Evolutionary Anthropology*, *5*, 172–180.
- Byrne, R. W. (1997). The technical intelligence hypothesis: an additional evolutionary stimulus to intelligence? In A. Whiten & R. W. Byrne (Eds.), *Machiavellian intelligence II: extensions and evaluations* (pp. 289–311). Cambridge: Cambridge University Press.
- Byrne, R. W. (1998a). Cognition in great apes. In A. D. Milner (Ed.), *Brain and cognition in monkeys, apes and man* (pp. 228–244). Oxford: Oxford University Press.
- Byrne, R. W. (1998b). Imitation: The contributions of priming and program-level copying. In S. Braten (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 228–244). Cambridge: Cambridge University Press.
- Byrne, R. W. (1999a). Cognition in great ape ecology. Skill-learning ability opens up foraging opportunities. *Symposia of the Zoological Society of London*, *72*, 333–350.
- Byrne, R. W. (1999b). Imitation without intentionality. Using string parsing to copy the organization of behaviour. *Animal Cognition*, *2*, 63–72.
- Byrne, R. W. (1999c). Object manipulation and skill organization in the complex food preparation of mountain gorillas. In S. T. Parker, R. W. Mitchell, & H. L. Miles (Eds.), *The mentality of gorillas and orangutans* (pp. 147–159). Cambridge: Cambridge University Press.
- Byrne, R. W. (2000). The evolution of primate cognition. *Cognitive Science*, *24*, 543–570.
- Byrne, R. W. (2001). Clever hands: The food processing skills of mountain gorillas. In M. M. Robbins, P. Sicotte, & K. J. Stewart (Eds.), *Mountain gorillas. Three decades of research at Karisoke* (pp. 293–313). Cambridge: Cambridge University Press.
- Byrne, R. W. (2002a). Imitation of complex novel actions: What does the evidence from animals mean? *Advances in the Study of Behavior*, *31*, 77–105.
- Byrne, R. W. (2002b). Seeing actions as hierarchically organized structures. Great ape manual skills. In A. Meltzoff & W. Prinz (Eds.), *The imitative mind: Development, evolution, and brain bases* (pp. 122–140). Cambridge: Cambridge University Press.
- Byrne, R. W. (2003). Imitation as behaviour parsing. *Philosophical Transactions of the Royal Society of London (B)*, *358*, 529–536.
- Byrne, R. W. (2004). The manual skills and cognition that lie behind hominid tool use. In A. E. Russon & D. R. Begun (Eds.), *Evolutionary origins of great ape intelligence* (pp. 31–44). Cambridge: Cambridge University Press.
- Byrne, R. W. (2005a). Detecting, understanding, and explaining animal imitation. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: from mirror neurons to memes* (pp. 255–282). Cambridge, MA: MIT Press.
- Byrne, R. W. (2005b). The maker not the tool: The cognitive significance of great ape manual skills. In V. Roux & B. Brill (Eds.), *Knapping stone. A uniquely hominid behaviour?* (pp. 159–169). Macdonald Institute Monographs.
- Byrne, R. W., & Byrne, J. M. E. (1993). Complex leaf-gathering skills of mountain gorillas (*Gorilla g. beringei*): Variability and standardization. *American Journal of Primatology*, *31*, 241–261.

- Byrne, R. W., & Corp, N. (2004). Neocortex size predicts deception rate in primates. *Proceedings of the Royal Society of London: Biology*, 271, 1693–1699.
- Byrne, R. W., Corp, N., & Byrne, J. M. E. (2001a). Estimating the complexity of animal behaviour: How mountain gorillas eat thistles. *Behaviour*, 138, 525–557.
- Byrne, R. W., Corp, N., & Byrne, J. M. E. (2001b). Manual dexterity in the gorilla: Bimanual and digit role differentiation in a natural task. *Animal Cognition*, 4, 347–361.
- Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences*, 21, 667–721.
- Byrne, R. W., & Stokes, E. J. (2002). Effects of manual disability on feeding skills in gorillas and chimpanzees: A cognitive analysis. *International Journal of Primatology*, 23, 539–554.
- Byrne, R. W., & Whiten, A. (1988). *Machiavellian Intelligence: Social expertise and the evolution of intellect in monkeys, apes and humans*. Oxford: Clarendon Press.
- Cartmill, E. A., & Byrne, R. W. (2007). Orangutans modify their gestural signalling according to their audience's comprehension. *Current Biology*, 17, 1345–1348.
- Catmur, C., Walsh, V., & Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Current Biology*, 17, 1527–1531.
- Corp, N., & Byrne, R. W. (2002a). Leaf processing of wild chimpanzees: Physically defended leaves reveal complex manual skills. *Ethology*, 108, 1–24.
- Corp, N., & Byrne, R. W. (2002b). The ontogeny of manual skill in wild chimpanzees: Evidence from feeding on the fruit of *Saba florida*. *Behaviour*, 139, 137–168.
- Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild chimpanzees produce group-specific calls: A case for vocal learning? *Ethology*, 110, 221–243.
- Custance, D., Whiten, A., & Fredman, T. (1999). Social learning of an artificial fruit task in capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, 113(1), 13–23.
- Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, 20, 469–493.
- Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, 6, 178–190.
- Elliott, J. M., & Connolly, K. J. (1984). A classification of manipulative hand movements. *Developmental Medicine & Child Neurology*, 26, 283–296.
- Fox, E., Sitompul, A., & Van Schaik, C. P. (1999). Intelligent tool use in wild Sumatran orangutans. In S. T. Parker, H. L. Miles, & R. W. Mitchell (Eds.), *The mentality of gorillas and orangutans* (pp. 99–116). Cambridge: Cambridge University Press.
- Galef, B. G. (1988). Imitation in animals: History, definitions, and interpretation of data from the psychological laboratory. In T. Zentall & B. G. Galef, Jr. (Eds.), *Social learning: Psychological and biological perspectives* (pp. 3–28). New York: Erlbaum.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119, 593–609.
- Gardner, R. A., Gardner, B. T., & Van Cantfort, T. E. (1989). *Teaching sign language to chimpanzees*. Albany, NY: SUNY Press.
- Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural communication of the gorilla (*Gorilla gorilla*): Repertoire, intentionality and possible origins. *Animal Cognition*, 12, 527–546.
- Genty, E., & Byrne, R. W. (2010). Why do gorillas make sequences of gestures? *Animal Cognition*, 13, 287–301.
- Gergely, G., Bekkering, H., & Kiraly, I. (2002). Rational imitation in preverbal infants. *Nature*, 415, 755.
- Goodall, J. (1986) *The chimpanzees of Gombe: Patterns of behaviour*. Cambridge, MA: Harvard University Press.
- Gouzoules, S., Gouzoules, H., & Marler, P. (1984). Rhesus monkey (*Macaca mulatta*) screams: representational signalling in the recruitment of agonistic aid. *Animal Behaviour*, 32(1), 182.
- Green, S. (1975). Dialects in Japanese monkeys: Vocal learning and cultural transmission of locale-specific vocal behaviour? *Zeitschrift für Tierpsychologie*, 38, 304–314.

- Heyes, C. M., & Ray, E. D. (2000). What is the significance of imitation in animals? *Advances in the Study of Behavior*, 29, 215–245.
- Hobaiter, C., & Byrne, R. W. (2010). The gestural repertoire of the wild chimpanzee. *Animal Cognition*. DOI 10.1007/s10071-011-0409-2.
- Hobaiter, C., & Byrne, R. W. (2011). Serial gesturing by wild chimpanzees: its nature and function for communication. *Animal Cognition*. DOI 10.1007/s10071-011-0416-3.
- Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition*, 8, 164–181.
- Humphrey, N. K. (1976) The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303–317). Cambridge: Cambridge University Press.
- Janik, V. M., & Slater, P. J. B. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, 26, 59–99.
- Jolly, A. (1966). Lemur social behaviour and primate intelligence. *Science*, 153, 501–506.
- Kuhl, P. K. (1982). Discrimination of speech by non-human animals – basic auditory sensitivities conducive to the perception of speech-sound categories. *Journal of the Acoustical Society of America*, 70, 340–349.
- Loucks, J., & Baldwin, D. (2009). Sources of information for discriminating dynamic human actions. *Cognition*, 111, 84–97.
- Marler, P., & Tenaza, R. (1977). Signalling behaviour of apes with special reference to vocalization. In T. Sebeok (Ed.), *How animals communicate* (pp. 965–1033). Bloomington, IN: Indiana University Press.
- Matsuzawa, T. (2001). Primate foundations of human intelligence: A view of tool use in nonhuman primates and fossil hominids. In T. Matsuzawa (Ed.), *Primate origins of human cognition and behaviour*. New York: Springer.
- Matsuzawa, T., & Yamakoshi, G. (Eds.). (1996). *Comparisons of chimpanzee material culture between Bossou and Nimba, West Africa*. Cambridge: Cambridge University Press.
- Mitani, J. C., Hasegawa, T., Gros-Louis, J., Marler, P., & Byrne, R. (1992). Dialects in wild chimpanzees? *American Journal of Primatology*, 27, 233–243.
- Napier, J. R. (1961). Prehensility and opposability in the hands of primates. *Symposia of the zoological society of London*, 5, 115–132.
- Perner, J., & Wimmer, H. (1985). ‘John thinks that Mary thinks that’: Attribution of second-order beliefs by 5- to 10-year old children. *Journal of Experimental Child Psychology*, 39, 437–471.
- Pika, S., Liebal, K., & Tomasello, M. (2003). Gestural communication in young gorillas (*Gorilla gorilla*): Gestural repertoire, learning, and use. *American Journal of Primatology*, 60, 95–111.
- Range, F., Viranyi, Z., & Huber, L. (2007). Selective imitation in domestic dogs. *Current Biology*, 17, 868–872.
- Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1996). Premotor cortex and the recognition of motor actions. *Brain Research*, 3, 131–141.
- Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (2002). From mirror neurons to imitation: facts and speculations. In A. Meltzoff & W. Prinz (Eds.), *The imitative mind: Development, evolution, and brain bases* (pp. 247–266). Cambridge: Cambridge University Press.
- Russon, A. E. (1998). The nature and evolution of intelligence in orangutans (*Pongo pygmaeus*). *Primates*, 39(4), 485–503.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., & Rumbaugh, D. M. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development*, 58, 1–252.
- Seyfarth, R. M., & Cheney, D. L. (1986). Vocal development in vervet monkeys. *Animal Behaviour*, 34, 1640–1658.
- Spence, K. W. (1937). Experimental studies of learning and higher mental processes in infra-human primates. *Psychological Bulletin*, 34, 806–850.

- Stokes, E. J., & Byrne, R. W. (2001). Cognitive capacities for behavioural flexibility in wild chimpanzees (*Pan troglodytes*): The effect of snare injury on complex manual food processing. *Animal Cognition*, 4, 11–28.
- Stokes, E. J., Quiatt, D., & Reynolds, V. (1999). Snare injuries to chimpanzees (*Pan troglodytes*) at 10 study sites in East and West Africa. *American Journal of Primatology*, 49, 104–105.
- Tanner, J. E. (1998). *Gestural communication in a group of zoo-living lowland gorillas*. Unpublished PhD thesis, University of St Andrews, St Andrews.
- Tanner, J. E., & Byrne, R. W. (1996). Representation of action through iconic gesture in a captive lowland gorilla. *Current Anthropology*, 37, 162–173.
- Tanner, J. E., & Byrne, R. W. (1999). The development of spontaneous gestural communication in a group of zoo-living lowland gorillas. In S. T. Parker, R. W. Mitchell, & H. L. Miles (Eds.), *The mentalities of gorillas and orangutans. Comparative perspectives* (pp. 211–239). Cambridge: Cambridge University Press.
- Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states – the question is which ones and to what extent. *Trends in Cognitive Sciences*, 7, 153–156.
- Tomasello, M., George, B., Kruger, A., Farrar, J., & Evans, E. (1985). The development of gestural communication in young chimpanzees. *Journal of Human Evolution*, 14, 175–186.
- Tomasello, M., Gust, D., & Frost, T. A. (1989). A longitudinal investigation of gestural communication in young chimpanzees. *Primates*, 30, 35–50.
- Visalberghi, E., & Fragaszy, D. M. (1990). Do monkeys ape? In S. T. Parker & K. R. Gibson (Eds.), *Language and intelligence in monkeys and apes* (pp. 247–273). Cambridge: Cambridge University Press.
- Wellman, H. M. (1990). *Children's theories of mind*. Bradford: MIT Press.
- Whiten, A. (1998). Imitation of the sequential structure of actions by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 112, 270–281.
- Whiten, A., & Byrne, R. W. (Eds.). (1997). *Machiavellian intelligence II: Extensions and evaluations*. Cambridge: Cambridge University Press.
- Whiten, A., Custance, D. M., Gomez, J.-C., Teixidor, P., & Bard, K. A. (1996). Imitative learning of artificial fruit processing in children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 110, 3–14.
- Zuberbuhler, K. (2002). A syntactic rule in forest monkey communication. *Animal Behaviour*, 63, 293–299.