GESTURAL COMMUNICATION IN ORANGUTANS (PONGO PYGMAEUS AND PONGO ABELII): A COGNITIVE APPROACH

Erica A. Cartmill

A Thesis Submitted for the Degree of PhD
at the
University of St. Andrews

2008

Please use this identifier to cite or link to this item:
http://hdl.handle.net/10023/634

Gestural communication in orangutans (*Pongo pygmaeus* and *Pongo abelii*): a cognitive approach

Erica A. Cartmill

Thesis submitted to the University of St Andrews
for the degree of Doctor of Philosophy

June 2008

I, Erica A. Cartmill, hereby certify that this thesis, which is approximately 63,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.
date .... signature of candidate ....

I was admitted as a research student in October, 2004 and as a candidate for the degree of Ph.D. in October, 2005; the higher study for which this is a record was carried out in the University of St Andrews between 2004 and 2008.
date .... signature of candidate ....

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D. in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.
date .... signature of supervisor ....

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration.
date .... signature of candidate ....

Acknowledgements

I would like to thank my parents, Kaye Brown and Matt Cartmill, for their tireless enthusiasm and lifelong support of my intellectual development and scientific pursuits. The culmination of this thesis is a testament to their inspiration and encouragement.

I am also grateful for the guidance, help, editing, and ideas I received from my supervisor, Richard Byrne, and colleagues at St Andrews. Without them, the development of this thesis would have been far less interesting and enjoyable.

Abstract

While most human language is expressed verbally, the gestures produced concurrent to speech provide additional information, help listeners interpret meaning, and provide insight into the cognitive processes of the speaker. Several theories have suggested that gesture played an important, possibly central, role in the evolution of language. Great apes have been shown to use gestures flexibly in different situations and to modify their gestures in response to changing contexts. However, it has not previously been determined whether ape gestures are defined by structural variables, carry meaning, are used to intentionally communicate specific information to others, or can be used strategically to overcome miscommunication.

To investigate these questions, I studied three captive populations of orangutans (*Pongo pygmaeus* and *P. abelii*) in European zoos for 10 months. Sixty-four different gestures, defined through similarities in structure and use, were included in the study after meeting strict criteria for intentional usage. More than half of the gesture types were found to coincide frequently with specific goals of signallers, and were accordingly identified as having meanings. Both structural and social variables were found to determine gesture meaning. The recipient's gaze in both the present and the past, and the recipient's apparent understanding of the signaller's gestures, affected the strategies orangutans employed in their attempts to communicate when confronted with different types of communicative failure (e.g. not seeing, ignoring, misunderstanding, or rejecting a gesture). Maternal influence affected the object-directed behaviour and gestures of infants, who shared more gestures with their mothers than with other females. These findings demonstrate that gesture can be used as a medium to investigate not only the communication but also the cognition of great apes, and indicate that orangutans are more sensitive to the perceptions and knowledge states of others than previously thought.

Table of Contents

CHAPTER 4: USE OF GESTURAL SEQUENCES 127

CHAPTER 5: UNDERSTANDING OTHER MINDS 165

Index of Tables

Index of Figures

Index of Appendices

Chapter 1: Introduction

> *Laughter, weeping, anger, rage, various shades of desire, of disappointment, and of greed, are partially vocalised, but there is slight indication of meaning as contrasted with feeling, and the observers were finally led to conclude that, as a means of intercommunication, vocalisation is less important and bodily attitude and gesture more important in the chimpanzee than in man.*
> (Rothmann & Teuber 1915, pp. 13-14)

1.1 The origin of language: evolutionary perspectives

1.1.1 Paleoanthropology

Much can be learned about how early hominins lived, fed, and died from the study of their dwellings, tools, artefacts, food remains, and fossils. We know that by 3.6-3.75 million years ago (m.y.a.) human ancestors could walk upright (Leakey and Hay 1979), and that by 2.6 m.y.a. they were making and using stone tools (Semaw 2000). Around one m.y.a. ancestral hominins were using fire (Brain and Sillen 1988); by 400 thousand years ago (k.y.a.) they hunted with spears (Thieme 1997); and by 75 k.y.a. they made, and presumably wore, ornamentation (d'Errico et al. 2005). Evidence for burial appears at 61 k.y.a. in Neanderthals, though its function remains the subject of some debate (Cartmill and Smith in press). From these concrete traces, we can extrapolate an evolutionary timeline for the emergence of many distinctive behavioural characteristics of humans, including complex tool manufacture, cooking, religion, agriculture, and art.

One central element of human evolution that remains within the realm of speculation is the origin of language. We may be able to guess, with some certainty, what our early ancestors looked like, ate, made, and did, but we have very little idea what they sounded like. Did they live in a world of virtual silence, punctuated only by involuntary cries of pain or fear? Did they speak fluently, if simply, using a predecessor of modern human languages? Did they gesture to one another, using a combination of conventionalized signs and pantomime to be understood? Most likely, none of these caricatures ever existed; however, the lack of negative evidence makes all theories about the earliest form of language difficult to refute.

In an attempt to substantiate theoretical models of the emergence of language, researchers have tried to link language to the appearance of other, more traceable, abilities. Tool manufacture, bipedalism, art, and organized hunting have all been claimed as hallmarks of the origin of language (Montagu 1976; Aiello et al. 1996; Knight 1998; Ambrose 2001). While each of these abilities signifies a manifest cognitive or behavioural advancement over the behaviour of extant great apes, none are conclusively linked to any communicative improvement. The lack of "fossilization" of language until the advent of writing, again relegates such theories to the realm of speculation.

Endocasts of fossil hominin skulls help to identify the surface physiology of our ancestors' brains and may shed light upon the development of regions of the brain that are associated with cognitive or communicative abilities in extant primates. Cranial endocasts of *Homo habilis* reveal an impression of Broca's area on the left hemisphere, suggesting that the neuroanatomical changes associated with language had taken place by 300 k.y.a. (Ambrose 2001). It is not wise to ascribe modern language abilities to *H. habilis* on the basis of this alone, however, as chimpanzees have a structure resembling Broca's area (Gannon et al. 1998; Sherwood et al. 2003), and certainly do not possess a full language. Endocasts may be our only source of data about the structure of ancestral hominin brains, but their usefulness in making inferences about cognitive abilities is limited by the lack of direct correlation between surface structure and function.

Examination of genetic correlates of language, through linking heritable genetic structure to expressed linguistic ability, provides more information about function than endocasts can. DNA analysis of living humans with heritable language impairment is a good way to identify genetic markers associated with language, and provides some indication of the outcome of missing or disrupted genes important in language production or comprehension. The standard mutation rate of mitochondrial DNA or of genes on the

Y-chromosome can be used to extrapolate timelines for significant changes in the genome and thus provides a way to date genetic mutations that occurred in the past. If the capacity for language emerged as a result of genetic mutation, these techniques will prove an invaluable tool in placing the onset of language. However, inferring that language was present in ancestral hominins through the presence of genetic markers is subject to the same criticism as inferences based on endocast impressions of brains, namely that presence of some of the physiological structures necessary for language does not necessarily mean that individuals of that era used language. Abilities or structures that are correlated with language in modern humans may be *necessary* but not *sufficient* conditions for language.

The discovery of one genetic marker necessary for language has helped narrow down the timeline for the presence of fully-developed human language. Examination of a human family with heritable language deficiencies led to the discovery of the importance of FOXP2, a gene linked to the capacity for speech production (Lai et al. 2001). A study dating significant changes in the gene suggested that FOXP2 has been present in its current form for around the last 200,000 years (Enard et al. 2002). Assuming that possession of a fully-functional copy of the gene is a prerequisite for language, this discovery narrowed the evolutionary timeline for spoken language to the last 200,000 years. Recent study of Neanderthal remains, however, has found that Neanderthals carried the human-variant of FOXP2 (Krause et al. 2007). This contradicts the timeline predicted by (Enard et al. 2002) and suggests that the human form of the gene emerged in the last common ancestor of humans and Neanderthals, at least 300 to 400 k.y.a. Though there is some fossil and genetic evidence for the physiological capacity for speech in humans from between 100 k.y.a. and 200 k.y.a. (or in hominins as early as 400 k.y.a.), this does not categorically signify the presence of language in hominins during this span of time.

1.1.2 Anthropology

Writing, the only indisputable physical evidence of ancient language, did not appear until 4,000 BC (Schmandt-Besserat 1978), tens (if not hundreds) of thousands of years after human language came into being in its present form. An attempt to correlate genetic similarity of modern human indigenous peoples with linguistic similarity of their languages suggests that language change accompanied geographic shift in early populations (Cavalli-Sforza et al. 1988). Comparison of these population shifts with archaeological data has helped to date genetic and linguistic splits, suggesting that the first major split in genetic and linguistic groups occurred in Africa ≥ 92 k.y.a (Cavalli-Sforza et al. 1988). Though this technique helps to determine the rate of language change and to estimate dates for many of the major divisions between groups, it is not useful in reconstructing the date or form of the earliest manifestations of language.

In the past, researchers looked to illiterate tribes with little material culture for evidence of "primitive" languages that may have represented earlier stages in the evolution of language (Humboldt 1999). These studies, however, revealed that, although highly variable in form, existing languages do not map neatly onto any scale of complexity that could be representative of the emergent stages in our evolutionary history (O'Grady et al. 1997). Cross-cultural linguistic comparisons also gave rise to theories of language universals, which hypothesised that humans are born with a basic cognitive structure of language onto which the surface expressions of local words and grammar develop (Chomsky 1965; Chomsky 1972; Jackendoff 2002). Theories of language universals and "deep grammar" have been replaced in recent years by those which view the human infant as a generalist learner, capable of forming rules from a paucity of linguistic input (Tomasello 2003; Kirby et al. 2007). If humans are simply very good generalist learners and have no language-specific cognitive mechanisms, then it is likely that human linguistic abilities were subject to natural selection and evolved slowly rather than in a "catastrophic mutation" that some have claimed was responsible for the intact emergence of full human grammar (Chomsky 1957; Bickerton 1990). If language is indeed the product of generalist learning (rather than generative grammar), then one would expect non-human primates to have simpler or less refined forms of both human cognitive and communicative abilities rather than categorically different systems.

1.1.3 Non-human primates

Examination of the fossil record and comparison of the languages of modern humans provide some insight into the emergence of the modern human mind and voice, but cannot reveal what cognitive and communicative capacities were present in ancestral hominins before the recent evolution of the *Homo* lineage. The most convincing speculations about the mental and social life of human ancestors come from identifying cognitive and communicative homologies between modern humans and living non-human primates. By comparing and contrasting the natural communication systems of other primates (as well as their ability to learn human linguistic systems), we can determine not only what properties of communication were present in our last common ancestors, but also how these abilities were likely to have been used—their function as well as their form.

The study of the cognition and communication of other living primates is mainly focused on the great apes because they represent the most recent divergence from the human lineage. Chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*) split from the human lineage most recently, approximately 6 m.y.a.. Gorillas (*Gorilla gorilla* and *Gorilla beringei*) diverged between 7 and 9 m.y.a. and orangutans (*Pongo pygmaeus* and *Pongo abelii*) between 10 and 12 m.y.a. (Stauffer et al. 2001). Our genetic similarity to the great apes corresponds to the duration of our shared ancestry with the various species. Great apes share between 93 (orangutans) and 99 percent (chimpanzees) of our genome (King and Wilson 1975; Wildman et al. 2003).

Great apes distinguish themselves from lesser apes and monkeys through their cognitive similarity as much as by their genetic proximity to humans. All of the great apes have demonstrated abilities to recognize themselves in mirrors (Walraven et al. 1995; Heyes 1998; Posada and Colell 2007), a behaviour often linked to early stages of self-awareness. In addition, all genera of great apes have been shown to make and use tools in captivity and to use tools in the wild (Ingmanson 1996; Inoue-Nakamura and Matsuzawa 1997; Boysen et al. 1999; Fox et al. 1999; Parker et al. 1999). The ability to manufacture tools has long been regarded as an essential part of the evolution of early human culture (see Ambrose 2001 for review). Finding rudimentary tool manufacture in all four great apes indicates that the ability to evaluate one object for its usefulness in obtaining another is at least 14 million years old. Though some species of monkeys (Ottoni and Mannu 2001), cetaceans (Krutzen et al. 2005), and birds (Hunt 1996; Hunt 2000; Tebbich and Bshary 2004) have been found to use (and in the case of corvids, manufacture) tools, these abilities are more likely to be analogous traits as they are found in only a few distantly-related genera.

In the realm of social understanding, apes have demonstrated advanced abilities to act according to the gaze of other individuals (Liebal et al. 2004; Bräuer et al. 2005), and have been observed (along with monkeys) to utilize tactical deception (Byrne and Whiten 1988; Whiten and Byrne 1988; Byrne and Whiten 1992). Gaze following and the more complex ability of visual perspective taking are both possible predecessors to a theory of mind, as they require at least a functional understanding that other individuals have visual access that may differ from one's own. Higher-order tactical deception (observed in all great apes but not conclusively in monkeys) requires an ability to attribute intention or mental states to other individuals and is probably indicative of possession of an advanced theory of mind. Taken together, these studies demonstrate the complex social mind of the great apes, one in which other individuals are treated as autonomous agents with different sets of perceptions. Though apes fall short of the abilities of modern humans in attributing mental states to others, they do show a remarkable sensitivity to social relationships and an ability to manipulate others to achieve their own goals. This social acuity was most likely present in the last common ape ancestor before orangutans diverged from the African ape lineage.

1.2 Great ape communication

1.2.1 Vocalisation

When it comes to vocal communication, primates are far less flexible in their behaviour than they are in situations of social manipulation. Though many non-human primates use vocalisations in seemingly complex ways to inform others about predators (Seyfarth et al. 1980; Seyfarth et al. 1980; Zuberbühler 2000; Zuberbühler 2001) or food sources (van Krunkelsven et al. 1996; Roush and Snowdon 2000), or to advertise to others when they are affiliating with a high-ranking social partner (Cheney and Seyfarth 1990; Slocombe and Zuberbühler 2005), the vocal signals themselves are relatively fixed. Non-human primates may learn to coordinate vocalisation with a group or partner (Geissmann 1999) or learn to identify appropriate contexts for performing a specific call (Hauser and Wrangham 1987; Cheney and Seyfarth 1990; Hauser et al. 1993), but they seem unable to learn completely new calls (Janik and Slater 1997). In addition, it has been argued that much of the observed contextual variation in calls is due to difference in emotional arousal (Hauser et al. 1993; Tomasello and Call 1997).

In the 18th century, Charles Darwin described the complexity of ape postures and facial expressions and noted their similarity to our own (Darwin 1955). Scientists in the early 20th century who had spent time observing wild or captive chimpanzees remarked that ape gestures seemed more flexible and less emotionally-linked than their vocalisations (Rothmann and Teuber 1915; Yerkes and Learned 1925). These early observations of the flexibility of ape gestures have been supported by a recent study comparing the use of both gestures and vocalisations in *Pan* (Pollick and de Waal 2007). This study found that while both *Pan paniscus* and *Pan troglodytes* use vocalisations and facial displays in highly similar ways, there is significantly more inter-species (and individual) variation in the contextual use of gesture. This finding led the authors to conclude that gestures are less tied to certain behavioural contexts and can thus be employed with greater flexibility than vocalisations and facial displays.

1.2.2 Taught language studies

The main limitation on the complexity of ape vocal communication appears to be structural restriction of the physiology of the vocal production anatomy rather than a general inability to modify communicative actions in response to social cues (Fitch 2000). The specific vocal limitation is demonstrated by the relative success of studies attempting to teach apes to speak versus communicating using non-vocal linguistic systems. The first attempts to teach language to an ape were made by the Hayes family in the 1950s (Hayes and Hayes 1951). The Hayeses acquired an infant chimpanzee, Viki, and raised her alongside their infant son, hoping that she would acquire language and other human attributes if given near-identical rearing to that of a human infant. They found that, although Viki outpaced the child in terms of physical development and coordination, she was unable to speak. After intensive training and shaping of her mouth and tongue manually, the Hayes managed to teach Viki to say four words: “mama,” “papa,” “cup,” and “up.” It is likely that even this extremely limited success was a product of facial rather than vocal learning as her words were unvocalised, more like whispers or coughs than spoken words (Hayes 1951).

Though unable to alter their vocal repertoire to any great extent, apes have shown great aptitude, flexibility, and ability to learn in their use of visual forms of communication. Whereas Viki was unable to learn to speak, chimpanzees, gorillas and orangutans have all been successful at learning manual signs in American Sign Language and using signs to communicate their requests or to answer simple questions posed by their human keepers (Gardner and Gardner 1969; Patterson and Linden 1981; Miles 1990). Even semi-wild (rehabilitant) orangutans have been successfully taught to use signs to request different food types (Shapiro and Galdikas 1999). Most of the apes that were taught to use a standardised human sign language in an enculturated environment reached a level of proficiency comparable with that of a 2-year-old child. Additionally, a few individuals were reported to use sign language to display humour or deception (Patterson 1980; Jensvold and Fouts 1993), and occasionally to invent or combine known signs to refer to novel ideas or objects (Patterson 1980; Patterson and Cohn 1990; Fouts and Mills 1997). Though these studies often started out by teaching signs through moulding the hands of their ape subjects, apes often progressed to spontaneously matching or imitating signs (Gardner and Gardner 1969). One of the most advanced ape signers, a chimpanzee named Washoe, signed not only to her human caretakers but also to her adopted son, who successfully learned 51 signs from her and other chimpanzees with no human assistance (Fouts et al. 1989). The ability of apes to spontaneously acquire signs from other apes, without having been encouraged to learn signs from humans, demonstrates the ability to learn both the form and context of communicative symbols through observing others and is a skill required for human language.

Apes have also demonstrated great abilities in acquiring and communicating with graphic symbols. In a project led by Sue Savage-Rumbaugh, bonobos and chimpanzees acquired a keyboard-based system of symbols through observation rather than direct shaping (Savage-Rumbaugh et al. 1998). Though the main modality of the taught language was not gestural, the apes also spontaneously gestured to refer to objects, places, or people, and at least one employed an ordering preference (gesture last) when combining gestures with graphic symbols (Greenfield and Savage-Rumbaugh 1990). The success of apes in using human signs as well as their spontaneous inclusion of gesture within other taught modalities demonstrate the extent to which apes can use gesture to convey meaning within a human constructed linguistic system.

1.2.3 Natural gesture

After the success of studies teaching great apes to use human signed languages, researchers began to investigate the natural gestures used by these species to communicate with conspecifics or to request things from human caretakers. A comprehensive survey of the gestural repertoires of all four great apes was performed at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. By using similar methods for identifying gestures in each of the species, the authors found that each species had a repertoire of 20-30 gestures (Pika et al. 2003; Pika et al. 2005; Liebal et al. 2006; Call and Tomasello 2007). All four species were found to use both tactile and visual gestures, and the African species also used auditory gestures such as clapping or banging objects. The animals surveyed in those studies appeared to use their gestures appropriately in response to the visual attention of their recipient, choosing visual signals more frequently when they could be seen and other modalities when a potential recipient was looking away. Additional experimentation by the same group found that apes would move in front of a human experimenter before using visual gestures to request food (Liebal et al. 2004). These studies suggest that apes have some understanding of the communicative nature of their gestures—that they must be perceptible to be successful. Research involving captive gorillas strengthened this conclusion by finding that a captive gorilla would cover her play face to hide it from a potential partner (Tanner and Byrne 1993). Taken together, these studies suggest that apes demonstrate many aspects of their social awareness and understanding of the minds of others through their gestural communication.

1.3 Human gesture

Though the gestures and postures of humans have been noted for centuries for their communicative potential in drama and oration (Quintilian 2006), the possibility that gesture could be as rich a communicative medium as speech was not seriously considered until the 20th century. Once researchers showed that signed languages had consistent patterns of handshapes and movement similar to the semantic and syntactic structures of spoken language, signed languages began to be considered full languages rather than versions of ritualized pantomime (see Sacks 1990). This acknowledgement of the potential linguistic complexity of gesture expanded the field of gesture research into the realm of linguistics and allowed for the possibility that gesture played a significant role in shaping human communication.

1.3.1 Gesture as speech

The use of gesture as the sole communicative medium comes easily to humans. Deaf children raised without a sign language model will create and use idiosyncratic signs that possess some of the properties of language including reference and displacement (Morford and Goldin-Meadow 1997; Goldin-Meadow 2005; Goldin-Meadow et al. 2005; Botha 2007). Furthermore, one study has suggested that groups of deaf children, given time, spontaneously produce a fully-functional sign language (Senghas et al. 2004). In this example, deaf children who were gathered together for the first time when a school for the deaf was established began to communicate amongst themselves using gestures even though the school encouraged only vocal communication. The system of gestural communication that was created by the first group of children closely resembled the gestures of the community, with meanings represented holistically through shape and movement. However, two “generations” of children later, the system of gestures used amongst students at the school had developed into a fully-functional sign language comprised of discrete combinable units.

Gesture adopts the properties and structures of language not only in hearing-impaired individuals, but also in hearing individuals when they are asked to communicate without words. Regular syntactic rules that are not characteristic of the gestures hearing individuals make while they talk (such as predictable argument structure) appear in their gestures when the same hearing individuals are asked to retell events using only gesture (Goldin-Meadow 2003). More strikingly, arguments are produced in the same order across cultures regardless of the argument structure exhibited by the local spoken language (Gershkoff-Stowe and Goldin-Medow 2002; Goldin-Meadow et al. in press). The structures of human language express themselves easily in a visual modality and do so spontaneously in both deaf and hearing individuals under the right circumstances.

1.3.2 Gesture accompanying speech

When humans speak they gesture. Some reports suggest that as much as 90% of descriptive utterances are accompanied by a gesture (Nobe 2000). Spontaneous gestures made by hearing individuals during speech have been shown to provide both usable information to the recipient and insight into the thought processes of the speaker (McNeill 1992; Alibali and Goldin-Meadow 1993; Goldin-Meadow et al. 1993). This suggests that human language, though overwhelmingly verbal, is usually a multi-modal communication system. While spoken language breaks up meanings into words and uses combinations to encode meaning, gesture allows for the transmission of holistic ideas, often transmitting information about the quality or nature of an event not encoded in the accompanying speech (McNeill 2000).

The information contained in a speaker's gesture may sometimes contradict that which is transmitted verbally. In these cases, gesture may be a more accurate indicator of underlying thought processes than speech. Studies testing children on Piagetian conservation tasks and mathematical equations found that children whose gestures contradicted their incorrect verbal answers (e.g. they pointed to each number but verbally skipped over some) learned to solve the problems faster than other children (Church and Goldin-Meadow 1986; Alibali and Goldin-Meadow 1993). The authors concluded that gestures may provide insight into non-verbal thought processes and reasoning, and that gesture-speech mismatch may indicate a readiness to learn before any measurable success is detected in verbal reasoning ability. Gesture accompanying speech may thus provide information both about ideas not contained in the speech stream, and the speaker's underlying thought processes; communicative content as well as mental organisation.

1.4 Gestural origins of language

From the ubiquity of gesture in modern human communication and the flexibility of gesture used by extant great apes, many authors have drawn the inference that language may have had a gestural rather than vocal origin (Hewes 1973; Kendon 1991; Armstrong et al. 1995; Corballis 2002). The absence of vocal learning and lack of language-like characteristics in the vocalisations of great apes is perhaps the strongest argument for a gestural origin of language. Though the gestures of apes do not seem to express the linguistic features characteristic of human language (notably displacement and syntax), taught sign language studies show that apes are able to learn new meaningful gestures and to use them flexibly to reference objects in their environment.

The discovery of mirror neurons (specialized neurons that fire for both the doing and watching of a manual action) in both humans and monkeys has helped to strengthen theories of gestural origins by providing a neural basis for gestural imitation and learning (Rizzolatti and Arbib 1998; Arbib 2002; Arbib 2005). Mirror neurons found in Broca's area in humans and area F5 in monkeys are activated when a monkey performs a grasping action or observes a similar grasping action made by another (Rizzolatti et al. 1996). It has been suggested that this simple "mirror system" in monkeys could have provided the foundation for imitation of actions in apes and eventually could have evolved into "proto-sign," "proto-speech" and then language in the human lineage (Arbib 2005). Though mirror neurons do not confer any linguistic ability, they may provide the foundation for the ability to recognise that another's actions are the same as one's own, a recognition that would be a prerequisite for developing a system of meaningful standardised gestures.

1.5 Intentionality and other minds

The identification of communicative gesture relies of determining which movements are intentionally communicative. All movement is capable of transmitting information; it is possible by observing an animal's walk to determine that it is lame or by observing its persistent scratching to see that it has parasites. Neither of these potentially informative movements could easily be considered gestures however, as they are not primarily communicative. It is relatively easy to determine when an individual is producing a vocalisation, though it may be difficult to determine the signaller's motives. In contrast, gestures share a modality with many of the functional non-communicative movements of daily living, and so discerning which movements are communicative depends to a great degree on identifying which ones are intended as signals. Vocalisations may be studied and catalogued without actively addressing whether they are intentional signals, but communicative gestures must be defined partially through their intent as communicative signals, and thus the hurdle of attributing intentionality must be overcome.

In the field of human gesture, spontaneous movement of the hands and head made during speech is counted as gesture even though it may not be intended to communicate anything. Human gesturers however, though often unaware of the nature of their manual movements, are already communicating intentionally through their speech stream, and so it is assumed that the multimodal signal of speech plus gesture is directed towards their interlocutor.

When studying non-human animals or extremely young children, it is not possible to rely on language to reveal either intent or meaning. In order to attempt to ascribe first-person mental states (knowledge, intentions, etc.) from a third-person perspective, one is forced to rely on external behavioural cues. These cues may be poor indicators of the complexity of internal states, but they should lead to an underestimate not an overestimate the cognitive abilities of the subject. Testable criteria for demonstrating intentionality in both actions and communicative signals typically require that the signaller has both a goal and a flexible path towards achieving it. This "goal plus flexible path" combination features heavily in psychological literature from fields as diverse as child development (Piaget 1952; Bruner 1981), comparative cognition (Liebal et al. 2006; Pika *in press*), and neuroscience (Dickinson and Balleine 1994). By assuring that an animal or infant has both a goal and some degree of flexibility in reaching it, the likelihood of labelling inflexible or "hardwired" patterns of behaviour as intentional decreases.

1.5.1 Defining intentional gesture

In attempting to determine which movements might be intentionally communicative gestures, it is necessary to first rule out those actions that achieve their goals directly. A *communicative* act must indicate a desire or intention to another without directly bringing about the outcome itself. By this definition, reaching towards a desired object could be a communicative act (provided that nothing were preventing the object from being seized), whereas reaching out and picking up the object could not. Many communicative acts are incomplete versions of the functional action they represent, such as tugging gently on another's arm to indicate a desire for him to follow rather than forcibly pulling him. Communicative acts may also have no direct relationship to their goal, clapping to gain the attention of another for example is not an incomplete version of an effective action such as turning the head of the other individual. In view of the requirement that communicative acts achieve their goals *indirectly*, gestures must be physically ineffective.

Though “gesture” commonly refers to manual action, the most widely used definition of gesture in non-human primates includes movements of the head, limbs, and body and encompasses visual, tactile, and auditory (but non-vocal) actions (Tomasello and Call 2007). All movements could potentially be gestures as long as they fulfil other requirements of intentional signals. Actions produced automatically in response to specific stimuli are considered non-intentional as they are produced without strategy or flexibility. To be an intentional communicative signal, an act must be 1) directed towards another with 2) the objective of obtaining a goal and 3) be employed flexibly rather than as an automatic response to a stimulus.

Directed towards an audience

When assessing the intentionality of communicative signals rather than non-communicative behaviour, it is necessary to establish that the signal has not only an intended outcome, but also an intended *recipient*. The first step towards discovering whether a signal is directed to another individual or group is to determine whether the behaviour is performed when the animal is alone or only in the presence of an audience. So-called “audience effects” have been brought forward in claiming that a particular behaviour is a signal directed to others (Marler et al. 1986; Karakashian et al. 1988; Geissmann 1999), but the presence of an audience could simply be another necessary element of the stimuli that elicit the behaviour, thus allowing the possibility that the behaviour is an automatic reaction to a particular context. Audience effects must therefore be an initial filter in identifying intentional signals, but must not be used to claim intentionality of a communicative behaviour without further evidence.

In order to be intentionally communicative, an act must be directed toward a specific audience with the goal of modifying the recipients’ behaviour (or knowledge) in some way. For a potential gesture to be considered as directed towards another individual, it must either be produced while the signaller is oriented towards the other individual, or be followed by some measure of expectation of response from another individual or group (such as waiting for a response). If an act were produced preferentially when it was detectable by another individual (e.g. a visual signal produced more often when others are looking), that fact would strengthen the argument for the act's inclusion as an intentionally communicative signal.

Employed flexibly to reach a goal

Intentional communication requires not only an intended recipient of the signal, but also a goal; it is necessary to communicate *to* someone *about* something. Goal-directed behaviour is identified through persistence when the supposed goal is not met and ceasing of the behaviour when the goal is obtained (Bruner 1981). For a social action to be deemed to be intentional, the signaller must exhibit some external measure of expectation that the action will alter the behaviour of the other (Tomasello and Call 2007). These measures of expectation may take the form of waiting for a response or persisting if the other's behaviour remains unaltered. In order to attribute an intended goal to an agent, persistence in attempting to reach a goal must be flexible (Bruner 1981). Assigning flexibility as a criterion for goal-directed behaviour avoids attributing intentionality to actions produced as automatic reflexes to the environment. These automatic actions may demonstrate audience effects (as previously noted), and might be used persistently (through repetition) if the environment did not change. It is unlikely, however, that automatic actions would meet the criteria of flexibility as they are linked to specific stimuli and would likely be.

Demonstrating that a gesture has a clear goal and follows a flexible path towards reaching that goal goes a long way towards demonstrating communicative intent on the part of the signaller. The "goal plus flexible path" criterion has been identified as a marker of intentionality in child development (Piaget 1952; Bruner 1981), comparative cognition (Liebal et al. 2006; Pika in press), and neuroscience (Dickinson and Balleine 1994).

Though it is impossible to be entirely certain that a movement performed by an ape is an intentionally communicative gesture, the criteria enumerated above would help to establish both a communicative goal and voluntary control over the actions employed to reach that goal. In order to be counted as gestures in this study, actions had to be movements of the limbs, face, or body that were 1) not physically efficacious, 2) directed towards another individual, and 3) employed flexibly to reach a communicative goal. The specific criteria used to attribute intentional status to orangutan gestures observed in this study are described fully in Chapter 2.

1.6 Acquisition of gestures

The method by which natural gestures develop in non-human primates remains unclear. Researchers attempting to teach American Sign Language to enculturated apes claim that their subject added new signs to their repertoires by both imitation and invention, and that at least one infant acquired his initial repertoire of signs solely from conspecifics (Gardner and Gardner 1969; Fouts et al. 1989). The emergence of natural gesture, however, has yet to be systematically documented. Without longitudinal study of gesture use in infant apes and detailed records of what gestures they have been exposed to, researchers can only speculate on how and when gestures emerge in young apes. One study chronicled the changes in individual chimpanzees' gestural repertoires over a four year period and concluded that most could not be learned through second-person or third-person imitation (Tomasello et al. 1989). However, the sampling approach employed only compared the presence or absence of different gestures from two different years, and the individuals' exposure to gestures and gesture use in the four years between sampling was not known. It is difficult therefore, to draw firm conclusions about the ontogeny of specific gestures.

There are several ways in which a movement could theoretically become a meaningful gesture in the repertoire of an ape; gestures may be inherited, learned, or ritualized. It is possible to speculate on the origin of a particular gesture by comparing its frequency and use between individuals in the same population and in other populations (Tomasello and Call 2007), or within a population at different points in time (Tomasello et al. 1989). Though direct observation of exposure to and practice of new gestures would be necessary to pinpoint the developmental mechanism for each gesture, each theory of gestural development predicts a slightly different distribution of the gesture in the repertoires of individuals across different populations.

1.6.1 Heredity

It is possible that ape gestures are largely innate, as some claim most primate vocalisations are (Tomasello and Call 2007). If this were the case, we would expect that rearing history would have little effect on the form of individual gestures or the overall repertoire of gestures expressed by each individual. Thus, repertoires of individuals would not vary much either within the same population or between different populations. If gestures are much the same as vocalisations, it may be the case that even though the form is primarily genetically fixed, individuals learn aspects of usage such as the appropriate context through observation or practice (see Seyfarth and Cheney 1997).

1.6.2 Cultural transmission

Gestures might be acquired from others through social learning, through either observation of exchanges between two other individuals or participation in exchanges initiated by a knowledgeable signaller. Observation of a third-party gestural exchange can provide information about both the form and the goal of the gesture. A learner's participation as a recipient in a gestural exchange, even if the learner is ignorant of the meaning of the movement, could provide the same information if the signaller were to persist by escalating his action into the physically-effective action used to obtain his goal.

In this way, a learner could associate the signaller's initial gesture with his final goal, though active participation in future exchanges would require the cognitive leap of understanding that the recipient is expected to fulfil the goal once the initial gesture has been produced. Regardless of the specific mechanism of social learning, if a gesture has been socially learned, one would expect its distribution to be limited to members of one group rather than spread across different populations.

1.6.3 Ontogenetic ritualisation

It has been suggested that apes acquire most of their gestures through a developmental process of "ritualising" effective actions directed towards other individuals (Tomasello 1996; Tomasello and Call 2007). Tomasello and Call (2007) describe this process of so-called "ontogenetic ritualisation" as consisting of the following steps:

- Individual A performs behaviour X (not a communicative signal);
- Individual B consistently reacts by doing Y;
- Subsequently B anticipates A's performance of X, on the basis of its initial step, by performing Y; and
- Subsequently, A anticipates B's anticipation and produces the initial step in a ritualized form (waiting for a response) in order *to* elicit Y.

This process could be a highly effective mechanism for creating physically ineffective or "ritualised" gestures from full actions, but would perhaps not be applicable to cases where the gesture does not resemble a movement associated with end goal (as in the case of clapping or performing a headstand).

It is important to remember that the process of ritualisation is primarily one of individual learning. There is no potential for social learning of the gestures of others, but only for associating the initial movement of an action with the signaller's desired outcome or the recipient's likely response. This means that every individual must learn to produce gestures and to respond to gestures separately with every other individual. Without a mechanism for adopting the gestural forms of others, it is highly unlikely that individual A's ritualised gesture for Y would resemble individual B's gesture for Y. The only situation in which we would expect the gestures of different individuals to resemble each other is in cases where the gesture very closely resembles the final action (such as reaching or play biting).

Tomasello and Call (2007) state that if the variability in gestural repertoires within a group is as large as that between groups, the gestures were likely developed through ontogenetic ritualisation. However, they claim that gestures develop through ritualisation if they cannot be shown to be culturally transmitted (i.e. variation between groups is no higher than within groups, providing no support for culture). The only alternative method of gestural acquisition to ritualisation they consider is cultural transmission. This seems overly simplistic as it does not consider other potential means of gestural development or take into account whether ritualisation is a likely explanation for a particular gesture. Unless the gestural repertoires they compared were comprised solely of gestures that closely resemble the actions used to reach their end goals, then it is doubtful that ritualisation could account for the overall similarity in form. If ontogenetic ritualisation is indeed the mechanism by which gestures develop, then we would expect to see similarity (both within and between groups) among gestures that resemble their associated actions, and high individual variation among gestures that do not.

1.7 This study

I conducted the present study in order to examine the natural communication of orangutans from a cognitive perspective. Gesture, as opposed to vocalisation, was selected as the medium of interest for this study as all species of apes have demonstrated flexible use of natural gestures in captivity (Pika et al. 2003; Pika et al. 2005; Liebal et al. 2006; Call and Tomasello 2007). There is also strong evidence that, in apes, gesture is used more flexibly than vocalisation (Pollick and de Waal 2007). Flexibility is especially important in investigating possible antecedents to human language as human language is infinitely applicable, and words are not limited by a one-to-one mapping to external referents (as context-specific primate signals such as alarm calls seem to be), but can instead be used in multiple contexts to describe events present, absent, and imagined (Hauser et al. 2002). Reference, a core property of language, can be achieved by methods other than direct semantic labelling. Deictic gestures in particular, such as pointing, are used to direct the attention of other individuals but display great flexibility of use as they derive their meaning through their contextual use (Gómez 2004). It is possible that some gestures of apes manage to transmit meaning while still retaining flexibility in their use. Additionally, the flexibility of the gestural medium suggests that gestures in apes may reflect underlying cognitive states or processes and be a better indicator of communicative ability or intent than vocalisation.

1.7.1 Choice of species

Though the great apes are often lumped together through their anatomical and cognitive similarities that differentiate them from monkeys and lesser apes, the African great apes are more closely related to humans than they are to orangutans (Ruvolo 1997; Wimmer et al. 2002). Orangutans diverged from the hominin lineage more than 10 m.y.a. (Stauffer et al. 2001), and are the best living representative of the last common ancestor of all the great apes and humans (Byrne 1995). Orangutans were chosen as the study species as they might be used to identify communicative or cognitive plesiomorphic or synapomorphic traits within the great ape lineage. If one is interested in reconstructing what cognitive and communicative abilities were present in Hominoidea long before any human-specific traits evolved, modern orangutans are an ideal species to investigate. If orangutans share linguistic or cognitive traits with humans, it is likely that these are homologous traits and are found in the other great apes. Abilities shared by *Pongo* and *Homo* can be assumed to represent ancestral states or abilities that are at least 10 million years old and long predate human-specific linguistic advances. Conversely, where orangutans can be shown to use simpler communicative structures or fail to demonstrate a level of social understanding that is found in humans and *Pan* or humans and all

African apes, we can assume that these traits are synapomorphies that have evolved in the last 10 million years.

Few studies have been conducted on the communication of orangutans (either *pygmaeus* or *abelii*). The semi-solitary fission-fusion social structure of wild orangutans (van Schaik 1999) does not lend itself easily to studies of visual signalling between adults, and most studies of vocalisation have been aimed at determining the function of the male long call (MacKinnon 1974; Galdikas 1983; Mitani 1985). Other vocalisations have been catalogued (MacKinnon 1974), but none studied in depth.

Gestures in wild or reintroduced orangutans have been catalogued (MacKinnon 1974; Rijksen 1978; Bard 1992); but, with the exception of Bard (1992), these studies have paid little attention to the function or development of the gestures. The most comprehensive study of the gestures of captive orangutans to date has been that of Katja Liebal (Liebal et al. 2006; Liebal 2007). Her study was part of an attempt by the Max Planck Institute for Evolutionary Anthropology in Leipzig to catalogue the gestures of all of the great ape species using the same criteria and methodological approaches (see Call and Tomasello 2007).

The present study moves beyond previous work cataloguing gestural repertoires of orangutans by exploring the cognitive abilities underlying orangutans' use and development of gesture. By constructing a coding system specifically for use on orangutan gestures, I was able to tailor many of the parameters of measured variables to the temperament of the genus (rather than applying a definition more appropriate to an African ape). Through attempting to identify strategies or structures in the gestural communication of orangutans that suggested the use of cognitive properties shared with humans, I hoped to gain insight into the probable use of gesture in our last common ancestor.

1.7.2 Aims

The apparent rigidity of non-human primate vocalisations when compared to their gestures have led many researchers to speculate that the origins of human language were gestural rather than vocal (Hewes 1973; Kendon 1991; Armstrong et al. 1995; Corballis 2002). In studying the gestural communication of a great ape that shared an ancestor with humans more than 10 million years ago, I hoped to identify gestural structures or communicative strategies that likely represented shared ancestral traits of all great apes and might shed light on whether gesture use in apes can inform us about the evolution of human linguistic abilities. Certain cognitive abilities that humans rely on when they use language (primarily theory of mind) do not seem to be used by great apes in their vocal communication (Tomasello and Call 1997). In the current study, I aimed to determine whether orangutans used abilities related to theory of mind (such as use of others' visual perception or understanding) in their gestural communication. To this end, I compared orangutans' use of sequences of gestures in different social situations, to recipients who could or could not see them, as well as to recipients who did or did not understand their gestures. By comparing the use of communicative strategies in these different social situations, I hoped to learn more about the social understanding of orangutans and discover whether gesture is an effective medium for identifying cognitive abilities in non-human primates.

Previous studies that focused on or included the natural gestures of orangutans served mainly to catalogue their gestures and identify repertoires for the species and particular groups (MacKinnon 1974; Rijksen 1978; Liebal et al. 2006; Liebal 2007). These studies provided an important foundation for identifying the types of gestures used by orangutans, but failed to adequately address whether gestures are used as intentional signals. In comparing aspects of orangutan gestural communication to human language, and in using communicative strategies to gain insight into the social cognition of the species, it is essential to focus on signals that are used to intentionally communicate with others. Through identifying gestures initially by their structural similarities and then subjecting them to strict criteria of intentional usage, I aimed to create a repertoire of intentional gestures that was based on each individual's use of gesture rather than on the presence or absence of types of gestures on the group- or species-level. By focusing on the gestures used by each individual as well as the choices signallers made when faced with different communicative challenges, I hoped to determine the degree to which the strategic use and development of orangutan gestures could inform us about the underlying cognitive processes employed in the communication of orangutans. The cognitive development of orangutans was further explored through studies of triangular interactions in mother-infant dyads and maternal influence on the gestural repertoires of infants. Taken as a whole, the lines of inquiry presented in this thesis provide insight into the social cognition, theory of mind, cognitive development, and communicative strategies of the genus *Pongo*. Through the investigation of the social mind of the orangutan, this study aims to shed light upon our own cognitive and communicative history.

Chapter 2: General methodology

2.1 Introduction

The methods of data collection and analysis presented in this chapter pertain to the study of orangutans' use of gesture with conspecifics (Chapters 3 & 4). Two other studies are included in this thesis: one on orangutans' ability to modify their requests to overcome miscommunication with a human experimenter (Chapter 5), and one on the influence of orangutan mothers on the behaviour and gestures of their infants (Chapter 6). The methods used for the two latter studies are described in their corresponding chapters.

In order to gain insight into the cognitive and communicative abilities of orangutans, I performed an observational study of the naturally-occurring gestures of captive orangutans. My goal was to let the behaviour of the animals indicate which movements were meaningful communicative signals rather than classifying movements as gestures by their similarity to human gestures. To accomplish this, I took a structural approach to the classification of movement and attempted to remain as objective as possible. I initially identified and grouped examples of movement types together by the direction, speed, and quality of movement, and then combined or divided sets of examples according to their structural variables to create the most homogenous and distinct groups possible. By taking this structural approach, I hoped to define gestures solely through their movement and avoid using the context or behaviour of the recipient in the definition of a gesture.

I also addressed the issue of intentionality, employing strict criteria for intentional usage of a gesture in an attempt to ensure that only intentional movements were included as gestures. The criteria were applied to all gestures in the sample, and if an individual was not observed using a particular gesture type in an intentional way, the gesture was not included in his repertoire. By adopting an approach to intentionality that was based on the individual, I was able to account for the fact that, though a movement might be common, some individuals might not use it as an intentional gesture.

To determine gesture meaning, I analysed gestures for which the recipient's response matched the presumed goal of the signaller, labelling these gestures as having "observable goals." A gesture was deemed to have a specific meaning if it was predictably used in conjunction with one observable goal. Ambiguous gestures (those used frequently to achieve more than one goal) were examined further to determine whether the inclusion of additional structural or social variables into their definitions could reduce their ambiguity. Sequences of gestures were analysed (Chapter 4) to determine whether they represented persistence of the signaller's initial goal were combinations of gestures with different meanings and might represent syntactic combination of signals. The analysis of sequences also addressed the strategies signallers used in overcoming miscommunication and correspondingly focused on the different types of persistence used by signallers in varying social contexts. These analyses therefore had to take into account not only the structural variables of the gesture, but also social variables pertaining to the signaller and recipient (e.g. who can see whom, what the relationship between the two is), the recipient's reaction, and the subsequent actions of the signaller.

2.2 Subjects

Twenty-eight orangutans were observed during the study. This sample was comprised of two groups of Bornean orangutans (*Pongo pygmaeus*) housed at Twycross Zoo, UK and Apenheul Primate Park, NDL, and one group of Sumatran orangutans (*Pongo abelii*) housed at the Durrell Wildlife Conservation Trust, English Channel Islands. All orangutans were housed socially in groups of between 2 and 9. All three collections were comprised of a dominant male and several adult females with infant or juvenile offspring. The composition of the three collections is reported in Table 1. All together, the sample consisted of 15 adults, 6 juveniles (3 to 11-years-old), and 7 infants (under 3-years-old). One individual, an adult male (“Tom”), was transferred to another zoo 4 weeks into the observation period and was thus not included in many of the analyses. Two infants were born into the collection at Apenheul between data collection periods there, but were included in all analyses as the second data collection period was used mainly to gather data on them. A third infant at Apenheul was born 2 months before the second observation period but no data was collected from this infant and she was not included in the study.

Table 1: AGES (AT START OF OBSERVATION PERIODS), GENEALOGY, AND REARING HISTORIES FOR ALL INDIVIDUALS.

Collection	Individual	Sex	Birth date	Age	Age class	Birth and rearing	Mother (in study)
Apenheul *(Pongo pygmaeus)*	Binti	F	11/12/2000	4/5	Juvenile	Captive, briefly hand-reared	Fin (Sandy adopted)
	Dayang	F	01/12/2005	-/10 mo	Infant	Captive	Fin (Sandy adopted)
	Fin	F	10/05/1986	18/20	Adult	Wild	
	Jos	F	15/09/1992	12/14	Adult	Captive	
	Karl	M	??/??/1960	45/46	Adult	Wild	
	Katja	F	17/05/1997	7/8	Juvenile	Captive	Radja
	Radja	F	??/??/1962	43/44	Adult	Wild	
	Samboja	F	09/06/2005	-/16 mo	Infant	Captive	Sandy
	Sandy	F	29/04/1982	22/23	Adult	Captive	
	Silvia	F	19/12/1965	39/40	Adult	Captive	
	Tom	M	13/03/1989	16	Adult	Captive	
	Willie	M	17/04/2002	3/4	Juvenile	Captive	Radja
Durrell *(Pongo abelii)*	Dagu	M	19/04/1985	21	Adult	Captive, hand-reared	
	Gempa	M	24/05/2005	14 mo	Infant	Captive	Mawar
	Gina	F	??/??/1964	42	Adult	Wild	
	Jaya	M	13/05/2004	24 mo	Infant	Captive	Gina
	Jiwa	M	17/10/1999	6	Juvenile	Captive	Mawar
	Julitta	F	29/04/1975	31	Adult	Captive	
	Mawar	F	16/03/1989	17	Adult	Captive	Gina
	Putri	F	16/05/2005	14 mo	Infant	Captive	Julitta
Twycross *(Pongo pygmaeus)*	Batu	M	25/05/1989	17	Adult	Captive	
	Kibriah	F	23/01/1977	29	Adult	Captive, hand-reared	
	Maliku	F	10/06/1994	11	Adult	Captive	Kibriah
	Miri	F	03/08/2004	24 mo	Infant	Captive	Maliku
	Satu	F	31/08/1999	6	Juvenile	Captive	Theodora
	Tamu	F	29/11/2004	21 mo	Infant	Captive	Theodora
	Theodora	F	05/05/1988	18	Adult	Captive	Kibriah
	Tiga	M	22/07/2001	5	Juvenile	Captive	Kibriah

All three zoos housed orangutan groups containing both adult males and females, as well as juveniles and infants. The age and sex of all individuals is reported in Table 1 and illustrated in Figure 1. From Figure 1 we can see that Apenheul had the highest ratio of adults to immatures and Durrell had the highest proportion of males. The ages and sex of the orangutans at a particular zoo might have a strong influence on the types of behaviour observed at that collection. A group with a higher proportion of juveniles might have a higher frequency of gestures used in play, whereas a group that is composed of more adults might be observed using a greater number of agonistic gestures.

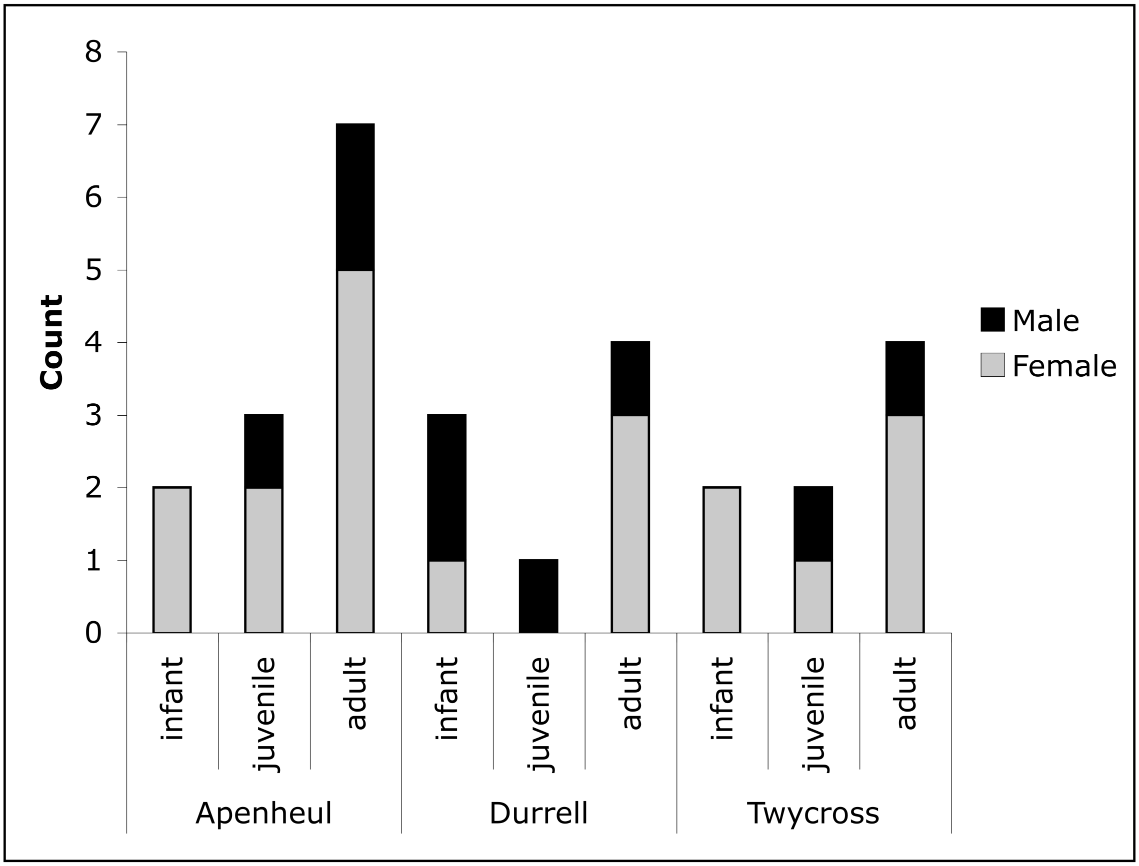


Figure 1: **AGE AND SEX COMPOSITION OF GROUPS**

2.2.1 Husbandry

Enclosures

All three groups were housed socially in enclosures with both indoor and outdoor areas. All outdoor areas contained large wooden or metal climbing structures, and those at Apenheul and Durrell included bushes and grasses that provided natural foraging opportunities. Photographs of both indoor and outdoor enclosures at the three zoos are included in Appendix I.

Apenheul Primate Park

The Apenheul enclosure was ring-shaped and could be divided into 4 sections, each connecting to an outdoor island. The 12 Bornean orangutans housed there could be split into 2, 3, or 4 subgroups throughout the day and all individuals except the two adult males and two oldest females were housed together at some point during the observation period.

Durrell Wildlife Conservation Trust

The 8 Sumatran orangutans housed at Durrell were kept in two indoor enclosures, each with access to its own outdoor island. They were split into two groups comprised on one side of a single matriline and an adult male and on the other of an adult female with her infant daughter and two white-handed gibbons (*Hylobates lar*). Though the orangutan dyad did not interact frequently with the gibbons, the adult female did appear to be highly vigilant whenever they were nearby, and their presence could have affected the type or quality of the interactions between mother and infant. Though separated, the two groups of orangutans had extensive contact through large metal mesh joining the two indoor enclosures. Adults interacted with each other by teasing, playing, or sharing food and infants frequently played with other individuals through the barrier.

Twycross Zoo

The group at Twycross was comprised of 8 Bornean orangutans housed in two indoor enclosures. Both enclosures had access to a shared outdoor area. The orangutans were split into two groups and were allowed access to the outdoor enclosure at alternating periods during the day. All orangutans at Twycross were part of a single matriline apart from the adult male. One group consisted of one female, her adult daughter and their two offspring. The other group was comprised of the same female's other adult daughter and her two offspring. The adult male and young female with her infant alternated between groups. The two groups had near-continual visual and auditory access to one another during the day.

Nutrition

During the periods of observation for this study, orangutans at all three zoos were given two main feeds of mostly whole fruits and vegetables (including oranges, bananas, grapes, apple, tomatoes, beets, leeks, carrots, celery, onions, cucumber, and lettuce) in the morning (between 7 and 10 AM) and in the afternoons (between 4 and 5 PM).

Apenheul Primate Park

Orangutans were given their morning and afternoon meals in separate enclosures. Older adults with no dependent offspring were fed in isolation; mothers and offspring were fed together. The standard diet of fruits and vegetables was supplemented with exotic fruits when available (such as durian fruit) as well as protein items such as chicken, eggs, cheese, nuts, and tofu. Apenheul provided whole vegetables, eggs, or cheese blocks as small feeds at least twice per day. These small feeds were given to individual orangutans at Apenheul by calling them down to the keeper access panels and handing them food items. Feeders with nuts were mounted inside and orangutans could obtain them by using tools to retrieve them. Water was freely available from indoor nozzles and from the outdoor moats.

Durrell Wildlife Conservation Trust

The adult male was separated off into a solitary enclosure during the morning and afternoon feeds. Other individuals were fed together from different feeders. Mother-infant dyads were occasionally separated from one another if competition or aggression was observed. A small fruit and vegetable meal was given on the outdoor islands at noon (by throwing a paper bag filled with food to each adult and juvenile). The vegetables and berries were mostly grown by the zoos own organic farm and provided freshly each day. Seasonal produce from the farm was supplemented (and in the winter replaced) with purchased fruit and vegetables. The farm produce was usually fed attached to the stem, eaves, or rest of the plat to provide a more natural foraging experience.

Small food items (such as nuts, seeds, berries, or monkey chow) were scattered onto the ground of the enclosure in the late morning or early afternoon by throwing small items of food into the undergrowth on the outdoor islands from across the moats. The outdoor islands had many species of grasses and bushes including several that produced edible berries. The orangutans at Durrell spent a large part of their time foraging in the plants on the islands, and some individuals foraged for water plants in the moats around the islands. Artificial termite mounds were filled with porridge approximately once a week, and honey was injected into holes in wall-mounted branches at the same frequency. Water was freely available throughout the day from nozzles on the wall, indoor pools, and the moat around the island.

Twycross Zoo

Orangutans were fed their two main meals in mother-offspring groups in their sleeping areas. The adult male was fed on his own. The standard fruit and vegetable diet was supplemented daily with bread, and several times per week with protein items (yogurt, tofu, eggs, or chicken). Small food items (grain, raisins, seeds) were hidden around the enclosure in the late morning or early afternoon by shutting the orangutans out and then scattering food items around the enclosure. Several times per week the orangutans were given short lengths of fire-hose into which jam or honey had been spread as enrichment. They used sticks to fish out the contents.

The outdoor enclosure at Twycross provided fewer opportunities for natural foraging than either Apenheul or Durrell (grass was the only plant growing there). All orangutans at Twycross were observed picking grass and eating it, some adult females "grazed" for most of their time outdoors. The keepers at Twycross attempted to provide more foraging opportunities by giving the orangutans freshly-cut branches from native English trees each afternoon. There was no source of fresh water at Twycross. The orangutans at Twycross were given bottles of diluted fruit drinks ("squash") or water with their daytime feeds and a keeper poured diluted fruit drink directly into their mouths during their last feed of the day.

Enrichment

As the observations took place at different times of the year in the three zoos, enrichment varied widely in response to whether the orangutans had access to their outdoor enclosures or not.

Apenheul Primate Park

My observations at Apenheul took place during the winter and early spring when the zoo was closed to visitors. The orangutans were closed inside for most of the time I was there (when the temperature was too cold) and so they were provided with many enrichment items that would not have been made available during the summer months. The keepers provided the orangutans with clothes and cloth of various strengths, which the orangutans made into hammocks, wore, and played with. Boxes and crates made out of cardboard and plastic were also provided. Occasionally, the orangutans were given brushes, which they used to brush their own hair or scrub against the walls. Paint and paper was provided to the adult male about once per month as he demonstrated an interest in drawing from time to time.

Durrell Wildlife Conservation Trust

At Durrell, no few artificial objects were provided as enrichment during the summer months when my observations took place. Artificial feeders built to elicit tool use were provided, but most of the foraging and object manipulation took place on the natural materials on the outdoor islands. The notable exception was that paper sacks were provided each day to each of the adults and juveniles so that they could use them as sunshades during the afternoon. Cardboard boxes were also occasionally provided and were used in nesting and object play.

Twycross Zoo

As the outdoor enclosure provided little natural enrichment, and only one of the two subgroups could access the outdoor area at one time, the orangutans at Twycross were provided with many objects indoors. They were given objects with pictures such as glossy magazines (which they flipped through looking at the pictures), stackable objects such as toilet paper tubes, and paper cups, objects that could make noise such as metal bin lids and plastic bottles, cardboard boxes of all sizes, and lots of cloth and clothes. Typically, all of these items were provided at once and the orangutans would make their way through the items during the day, often manipulating one type of object for as much as an hour before moving on.

Human Interaction

The levels of keeper and public interaction with the orangutans differed drastically between the three zoos. The husbandry routines and human traffic in the zoos could have affected the nature of the intra-group social interactions and the behavioural patterns of the orangutans in the different collections.

Apenheul Primate Park

My observations at Apenheul took place mainly when the zoo was closed to the public. There were four keepers whom the orangutans saw on a weekly basis, a few technicians, occasional guests of the zoo staff, and myself. The orangutans likely saw no more than 10 people per week. One of my observation periods lasted until the zoo opened for visitors. During this period, in the early spring, small family groups would pass through the orangutan exhibit every 10-20 minutes. I usually saw no more than 100 people during the week and 200-300 on the weekends. The visitors were mostly quiet and spent several minutes looking into all of the enclosures and commenting on the animals. One or two orangutans would approach the glass and watch the people, but most paid little to no attention. Importantly, the visitor area was about 3 meters above the ground level in the enclosures. There was a wide ledge that ran all the way around the edges of the enclosures at the same height as the visitor viewing area, but the orangutans could climb down out of view of the public. The glass separating the visitors from the orangutan was also broken up with coloured opaque sections, which decreased the visual access between humans and apes.

Durrell Wildlife Conservation Trust

Three keepers looked after the orangutans at Durrell, but one of them accounted for 70-80% of the days I was there. The zoo was open to the public during the entire observation period, but in the early morning and late afternoon there were typically no visitors. The orangutans tended to be fairly active during these periods and so many of my recordings were taken when there were no other visitors present. During the day there were always some people watching the orangutans, but the enclosures were designed so that there were many places where the orangutans could remain out of view. Unless the keepers were feeding, I could never see all orangutans at once, and there were many times when I could not see any. The islands were not flat and had a great deal of undergrowth. The visitor viewing areas across the moats were planted with trees and large grasses as well so the view from any one spot was always partially occluded. Most visitors had a hard time spotting the orangutans and many stayed for several minutes trying to see if they could spot one. The only time that the animals were predictably in full view was when the keepers few them, and so the 12 PM feed was accompanied by an outdoor talk given by the education department. This was the busiest time of the day and there could be up to 50 people gathered on the paths listening to the talk.

If it was raining and the orangutans stayed indoors, then anywhere from 2-30 people could be watching them through the glass at any one point. The number of people watching inside, even in these condensed conditions, was typically about two family groups at once (between 4 and 8 people) during most of the day. I witnessed few cases of people banging on the glass, and the orangutans almost always avoided the glass and ignored the visitors. The indoor enclosure was both tall and deep and had only one sitting platform near the visitors' side, so the orangutans only rarely came within 2 or 3 meters of the glass.

Twycross Zoo

The orangutans at Twycross were cared for primarily by a single keeper. A second keeper substituted for her on her days off or when she was ill. Nevertheless, this meant that 90-95% of the orangutans' caretaking routine was carried out by a single individual. The keeper also volunteered at an orangutan sanctuary in Borneo, and was particularly interested in the temperaments of the orangutans in her care. She provided the initiative for all of the enrichment routines, and had made significant progress in controlling obesity in the orangutans since she took over the section. Most of the decisions that governed the orangutans' lives were made or implemented by this one keeper.

According to its promotional materials, Twycross Zoo, has over 450,000 visitors annually; many of them school groups. The orangutan exhibit at Twycross experienced far more foot traffic than those at either of the other two zoos. During opening hours, there were at least 3-4 family groups or a school group observing the animals at any one time. Many of the visitors, particularly school groups would attempt to interact with the orangutans through the glass. Sometimes the orangutans ignored the visitors, but one female in particular, Theodora, spent stretches of time sitting at the window tapping on the glass or pulling faces when children approached.

The enclosure at Twycross provided no places where the orangutans could be completely out of the view of the public. The outdoor enclosure had a climbing structure but no occluding screens or natural foliage. The indoor enclosure had climbing structures near to the public area, which meant that the orangutans were often less than a meter from people watching them. Two of the three sleeping areas were adjacent to the public area as well, so when the orangutans had been shut into their sleeping areas at the end of the day, most were less than a meter away from the late afternoon visitors.

Potential husbandry effects

The enclosures, husbandry, and levels of visitor interaction varied drastically between the three groups. It is likely that these factors had a significant effect on the personalities, behaviour, and routines of the orangutans at each zoo. The orangutans at Durrell had the most naturalistic enclosure with many areas where they could retreat from view. The keepers at Apenheul changed the composition of the orangutan groups often so that the orangutans did not constantly have the same social partners. The observations also took place at Apenheul mainly when the zoo was closed to visitors. The orangutans at Twycross had less space and were more exposed to visitors than those at the other zoos. I also observed more stereotypic behaviour at Twycross (such as rocking, regurgitation, and licking the glass). Commonalities in behaviour and gesture between populations despite these differences are therefore more likely to be products of the temperament of the species than they are products of specific husbandry practices.

2.3 Apparatus and analysis

A Sony Handicam DCR TRV-38 was used to film the orangutans in both their indoor and outdoor enclosures. The miniDV videotapes were digitized onto a Macintosh Powerbook laptop and cut into clips using iMovie. Clips were coded using a spreadsheet designed in FileMaker Pro. The spreadsheet was sorted and analysed in Microsoft Excel. Paper coding sheets designed by the experimenter were used in one study to record the behaviour of mother-infant dyads. An Olympus digital voice recorder (VN-1000) was used in the same study to record continuous observations of the mother-infant dyads. The majority of statistical tests were carried out using Microsoft Excel, additional tests were performed using SPSS. All tests were two-tailed unless stated otherwise.

2.4 Data collection

Ad-libitum sampling was employed to gather video of social interactions likely to involve gesture (i.e. active social interactions such as locomoting, playing, object manipulating, or foraging). A pilot study suggested that focal-animal sampling, such as that employed by Liebal et al. (2006), would significantly reduce the number of gestures observed as active social interaction appeared to exist primarily in punctuated bursts.

Video was collected at Apenheul from March to April 2005 and in November 2006, at Twycross from April to May as well as in October 2006, and at Durrell from June to August 2006. This schedule resulted in approximately three months of observation at each of the three collections. Orangutans were observed for between 5 and 7 hours daily from areas accessible to the public. Video was recorded mainly during periods of the day when the potential for social interaction was highest (i.e. before and after feeding, and early or late in the day). Periods of rest and solitary feeding and solitary play were avoided. Extracted video clips of these periods of social interaction yielded 572 potential gestures from Twycross, 421 from Durrell, and 570 from Apenheul. The husbandry routines of the animals were not disrupted for filming. The observation period typically began in the morning between 8 and 10 AM after the animals had been fed and released from their sleeping enclosures and ceased between 4 and 6 PM when they were returned to them at the end of the day.

2.5 Video coding

Clips were extracted from the video and framed to contain a single social interaction including the context prior to the gesture and the recipient's reaction following. The social interactions of all individuals in the frame of the video were coded from the video. Though the video often included interactions involving more than two individuals, only dyadic interactions were included for analysis in this study. Clips ranged in length from less than 5 seconds to more than 3 minutes. Once digitised, the clips were coded using a self-designed template in FileMaker Pro (Appendix II).

2.5.1 Identification of potential gestures

Movements of the head, limbs, or body were identified as potential gestures and coded. The only initial criteria for inclusion were that the movements were "motorically-ineffective" and occurred in the presence of another individual. It should be noted though that, from the start of the study, I focused on those actions that were directed towards other individuals, and ignored some gestures performed away from others. This meant that my initial sample was likely biased towards intentional gestures as I selectively recorded directed gestures. Twenty-nine variables were coded for each potential gesture in the first round of coding, including the 1) animals involved in the interaction, 2) the specific form of the movement, 3) the response of the recipient, and 4) any attempts at persistence. A full list of the variables recorded and their definitions is provided in Table 2. The variables were chosen to allow analysis of the structural aspects of and social influences on gestures. To analyse structure, information about the limb or object used, the modality of the gesture, the facial expression of the signaller, and the target location of the gesture was recorded. Detailed information about the sequences of gestures and dialogues (including the recipients' response and signallers' persistence) was gathered so that interactive exchanges could be analyzed.

Coding of gestures ceased if an activity began that involved continuous contact or manipulative actions rather than gestures. Thus coding stopped at the onset of physical play, cuddling, nursing, carrying of an infant, or mating. If continuously interactive activities, such as those mentioned above, ceased for a least 10 seconds, any gestures following the pause were coded and included in the analysis.

Table 2: DEFINITIONS OF THE DIFFERENT VARIABLES CODED.

Examples of different possible values are given for each variable.

Type of variable	Variable coded	Definition	Examples
Social factors	Signaller	Individual performing a gesture	Gina, Kibriah, Sandy
	Recipient	Individual towards whom a visual gesture is oriented or who is touched during a tactile gesture	Gina, Kibriah, Sandy
	Relationship between signaller and recipient	Social/dominance relationship of signaller to recipient	Infant to mother, dominant to subordinate
	Context	Social context as defined by the most active interaction the signaller or recipient is engaged in with another individual or the environment	Play, affiliation, grooming, foraging, object manipulation
	Signaller visual attention	Whether or not the signaller can potentially see the recipient at the start of the gesture	Looking, not looking
	Recipient visual attention	Whether or not the recipient can potentially see the signaller at the start of the gesture	Looking, not looking
	State of recipient	Activity in which the recipient is engaged	Resting, watching, attempting play, carried, object manipulating
	Directedness of gesture	Oriented towards or touching another individual	Yes or no
Context within exchange	Role in dialogue	Assigned rank based on who initiated interaction	A or B
	Number of gestures in dialogue	Total number of gestures used by both parties in exchange	4
	Position of gesture in dialogue	Rank of current gesture in order of total gestures used in exchange	1
	Number of gestures in sequence	Number of gestures in continuous sequence (breaks no longer than 10 sec) by one individual	3
	Position of gesture in sequence	Rank of current gesture in order of total gestures used in sequence	2
Structure of gesture	Gesture	Type of gesture used (includes facial expressions made without limb movement)	Push, touch, wave, hit ground, play face
	Modality	Modality of the gesture	Visual, tactile, auditory
	Facial expression	Facial expression during the gesture	Play face, duck lips, teeth bared, lip pout
	Limb	Part of the body involved in the gesture	Right hand, left foot, head, whole body
	Handshape	Shape if the hand or foot used in gesture (if used)	Open flat hand, fist, finger
	Object	Object used to perform gesture	Foraged plant, stick
	Recipient body part	Part of the recipient's body towards which a movement is aimed or that is touched	Head, hand, mouth, torso

Table 2 continued:

Type of variable	Variable coded	Definition	Examples
Response and persistence	Response waiting	Signaller pauses after gesture, remains oriented towards recipient for at least 5 seconds	Yes or no
	Response	First reaction of supposed recipient within 10 seconds of gesture	Look towards, social play, hit, grab, leave
	Persistence	Signaller performs another gesture within 10 seconds of first	Yes or no
	Type of persistence	Relationship of second gesture performed to first	Repeat, exaggerate, change modality
	Gesture 2	Next gesture performed in sequence	Same as "gesture"
	Gesture 2 limb used	Part of the body involved in the second gesture	Right hand, left foot, head, whole body
	Gesture 3	Third gesture performed in sequence	Same as "gesture"
	Gesture 4	Fourth gesture performed in sequence	Same as "gesture"
	Result of exchange	Final result of interaction	Social play, recipient leaves, food sharing

2.5.2 Description of potential gestures

The potential gestures I observed took many forms; they could be based around a held position or a movement, use limbs or the whole body, be performed visually or contact another individual, and were often combined with other movements, facial expressions or body positions. To identify which social (as opposed to solitary) movements carried meaning for the orangutans in our study, I initially grouped movements that were similar in form together into potential gestures. I categorized gestures according to their modalities, location on body, movement pattern, force, and speed (Table 3). I assigned English glosses (i.e. "embrace") to each potential gesture based on its similarity to a human movement. Table 3 includes the number of gestures observed across all individuals. Individual use (of only intentional gestures) is given in Appendix IV.

Table 3: STRUCTURES AND DEFINITIONS OF ALL POTENTIAL GESTURES.

Seventy-one movements were initially identified as potential gestures. Here they are ordered within modalities by their value in each successive category (left to right). Categories appear in the order of: modality, location of the movement on the body, direction of motion, force, speed, and potential for use with an object. In the "object" category, "Poss" is an abbreviation for "possible" and an empty cell indicates that that movement could not be performed with an object.

Modality	**Body Part**	**Motion**	**Force**	**Speed**	**Object**	**Potential Gesture**	**Count**	**Definition**
Auditory	**Head**	Held	Any	Any		*Raspberry*	3	Lips are pursed and air pushed through them creating an audible buzz
	Hand(s)	Near to body	Any	Any		*Clap*	5	Hands are brought together quickly striking one another
		Down or away	Any	Any	Poss	*Hit ground/object*	50	Hand strikes the ground or an object (no weight is transferred)
Tactile	**Head**	Away	Low	Slow		*Kiss*	7	Closed, pursed lips are pressed to recipient's body
		Any	Low	Slow-Med		*Mouth*	11	Lips encircle or investigate recipient's body part (no teeth)
		Held	Any	Any		*Bite*	38	Open mouth is pressed against recipient and then partially-closed
	Hand(s)	Away	High	Fast	Poss	*Hit*	79	Hand touches recipient quickly and with force
						Grab	83	Hand closes quickly on recipient's body part
			High	Any	Poss	*Simultaneous hit*	9	Signaller and recipient "*hit*" each other at the same time (usually preceded by a slow coordinated arm raise)
			Med - High	Any	Poss	*Push*	63	Hand grips or rests on recipient and moderate or hard force is applied away from signaller
			Low	Slow		*Grasp*	87	Hand closes slowly on recipient's body part
				Slow-Med		*Investigate*	2	Gently explores a part of the recipient's body with fingers
					Poss	*Touch*	131	Hand gently comes in contact with recipient

Table 3 continued (part 2)

Modality	Body Part	Motion	Force	Speed	Object	Action	Count	Definition
Tactile	**Hand(s)**	Away	Low	Med	Poss	*Tap*	23	Fingertips lightly and quickly "hit" recipient
			Any	Med -Fast	Poss	*Poke*	5	Hand or finger exerts sharp directional pressure on recipient
				Any		*Pull away*	15	Hand closes on recipient and extends away from signaller
		Towards	Any	Slow		*Embrace/ pull*	7	Arm encircles recipient's shoulders or torso and "*pull*" is exerted
				Slow - Med		*Turn head*	25	Hand cups recipient's chin or cheek and rotates recipient's face toward signaller
				Any		*Pull*	148	Hand closes on recipient and retracts towards signaller
						Pull hair	37	Hand closes on recipient's hair and retracts towards signaller
					Poss	*Brush*	28	Hand lightly drags along recipient
		Held	Med - High	Slow		*Restrain*	4	Hand grips recipient and restricts recipient's movements
			Low	Slow		*Hold hand*	9	Hand closes gently around recipient's hand and remains there
			Any	Slow		*Embrace*	29	Arm encircles recipient's shoulders or torso
	Whole body	Away	Low - Med	Slow - Med		*Nudge*	34	Body part is placed in contact with recipient and directional force is applied briefly
		Up	Med	Any		*Shrug*	7	Shoulder is raised quickly against recipient
		Held	Low - Med	Any		*Tandem walk*	20	Hand placed on recipient's back during side-by-side walk

Table 3 continued (part 3)

Modality	Body Part	Motion	Force	Speed	Object	Action	Count	Definition
Visual	**Head**	Away	High	Fast		*Air bite*	14	Mouth opens and closes quickly as head moves towards recipient
			Low - Med	Slow		*Food beg*	15	Lower lip extended into scooped shape as head stretches forward or up
		Up	Any	Any		*Chin up/Nod*	2	Chin is raised and head tilts back or chin is lowered toward chest before returning to neutral position
		Held	High	Fast		*Teeth bared*	4	Lips retracted fully revealing parted teeth
				Slow - Med		*Duck lips open*	12	Lips are fully extended and flexed outward to create a trumpet shape (mouth is partially open)
						Top lip out	2	Upper lip is held rigid and extended forward and down partially covering mouth
				Any		*Raspberry face*	5	Mouth forms shape of raspberry vocalisation but without vocalising, movement is often repeated several times
			Med - High	Slow - Med		*Duck lips closed*	7	Lips are pursed together, fully extended and flexed outward to create a trumpet shape
						Whistle face	2	Lips are shaped into an "o" shape and extended forward
				Any		*Tense upper lip*	3	Upper lip extended down over teeth and held taught
			Low - Med	Slow -med		*Frog lips*	3	Lips flattened and broadened into a straight horizontal line
			Low	Slow - Med		*Look back*	8	Look back over shoulder at recipient while standing oriented away from recipient
			Any	Med -Fast		*Play face*	8	Corners of mouth are pulled back, mouth is opened wide and teeth are shown
				Slow - Med		*Lip smile*	9	Corners of mouth drawn back with lips closed
						Smile with teeth	5	Corners of mouth drawn back with lips parted, exposing closed teeth

Table 3 continued (part 4)

Modality	Body Part	Motion	Force	Speed	Object	Action	Count	Definition
Visual	**Head**	Held	Any	Any		*Pout*	5	Lips are pursed together and extended (as is “*kiss*”)
						Tongue out	4	Tongue extended past lips
		Held & directed	Low - Med	Slow		*Peer*	28	Torso angles towards recipient and head inclines towards recipient’s face or hands (position is held)
	Hand(s)	Away	High	Fast	Poss	*Fake*	9	Hand very quickly extended towards recipient and retracted
			Med - High	Fast		*Air grab*	16	Hand opens and closes quickly as arm extends towards recipient
					Poss	*Swat*	41	Arm swung quickly in arc towards recipient, palm facing recipient
			Any	Med-Fast		*Shoo*	19	Back of hand facing recipient, arm or hand extended quickly in arc towards recipient
		Away & held	Low - Med	Slow-Med	Yes	*Offer*	11	Object is extended towards recipient
			Any	Any		*Reach*	50	Arm and hand are both extended towards recipient
		Up	Med - High	Any		*Long body scratch*	2	Both arms scratch sides of body upward in exaggerated movement
			Any	Med-Fast	Poss	*Arms up*	21	One or two arms are extended and raised fully above head
		Up & held	Any	Any		*Raise arm*	29	Arm is extended from body and raised on the vertical axis
		Towards	Med - High	Any	Yes	*Drag object*	4	Object is pulled toward signaller (sometimes during locomotion)
			Low - Med	Any		*Cover*	8	Hand placed on top of object or part of signaller’s body
			Any	Any		*Beckon*	4	Arm, hand, or finger is extended toward recipient and swept in an arc toward recipient

Table 3 continued (part 5)

Modality	**Body Part**	**Motion**	**Force**	**Speed**	**Object**	**Action**	**Count**	**Definition**
Visual	**Hand(s)**	Towards	Low - Med	Slow-Med		*Wipe face*	2	Hand is cupped and brushed across mouth and nose
		Away & held	Low - Med	Slow	Yes	*Show*	4	Arm holding object is partially extended toward recipient and held
		Any	Med - High	Med-Fast	Yes	*Shake object*	24	Object held away from body and oscillated
			Any	Med-Fast		*Wave*	39	Arm held out from body and arm or hand oscillated
		Held	Any	Any	Yes	*Put object on head*	8	Object (usually cloth or paper) placed over signaller's head obscuring vision
	Whole body	Away	Any	Any		*Somersault*	63	Body curls forward into ball and rolls forward 360 degrees
			Any	Any		*Back roll*	12	Body curls into ball and rolls backward 360 degrees
		Away & held	Low - Med	Slow-Med		*Present body part*	4	Part of the torso or upper part of the limb is extended or angled towards recipient and held
		Any	Low - Med	Slow-Med		*Present genitals*	9	Legs are spread and genital region extended or angled towards recipient (facing recipient)
			Any	Any		*Roll on back*	15	Signaller moves to a position on her back (usually after performing a partial "*somersault*" or "*back roll*")
			High	Any	Poss	*Dangle*	29	Signaller hangs upside down from structure
			Any	Any		*Rock*	4	Body sways back and forth while on ground
			Any	Any	Poss	*Swing*	16	Body suspended from structure and moved through an arc (often repeated)
		Held	Med	Slow		*Headstand*	11	Torso inverted on ground so that pelvis is raised higher than head (often against a wall)

The categorisation of potential gestures in Table 3 was intended to objectively classify types of movement according to their structural characteristics. Each potential gesture has a different configuration of structural variables, and can be differentiated from other gestures by at least one of the variables listed in Table 3. This categorisation ensures that each gesture is structurally distinct and is able to be defined according to objective physical qualities. Most gestures are differentiated from one another by their definitions involving the variables displayed in Table 3, but some were differentiated by a variable that was not included on the table. If two gestures share the same modality, location, movement direction, force, speed, and object use, then the description in the far right column in Table 3 highlights the salient differences. This is especially important for facial gestures, as they all share the same modality and location, and often differ in only more specific qualities of movement (such as whether the teeth are visible or not). Manual gestures that had the same values for the variables in Table 3 most often were further defined by handshape or a specific quality or path of movement. Though presented hierarchically, all configurations of variables were possible, and variables were not ranked in subsequent analyses. The analysis of meaning presented in Chapter 3 helped to determine whether the initial categorisation accurately separated types of movement into gestures with unique structures and meanings, and helped me to redefine the definitions of some gestures.

In Table 3, I have first ordered the potential gestures by modality (e.g. whether or not they make a noise or come into contact with the recipient) as it seemed to be the most salient variable from the recipient's standpoint and had the greatest impact on the ways in which the recipient could perceive the gesture. I then categorised movements within the same modality by the part of body used. The body part used to make the gesture was defined broadly as head, hands/feet, or whole body. Movements sharing both a modality and a location on the signaller's body were further differentiated by the direction of movement. Motion direction was categorised in reference to the signaller's body axis. Movements were listed as either towards or away, up or down, performed near to the body, or held (stationary). I also used the force and speed of action to order movements. I defined force as either the pressure exerted on an object or the recipient in tactile gestures (low=soft, high=hard), or the energy (low, high) required to perform the action. In Table 3, the term "any" was used when the movement could vary in intensity from occurrence to occurrence but did not occur frequently enough or differently enough to warrant splitting it into more than one gestures. Both force and speed are subjective measures, and vary according to signaller, as a strong hit from an infant is different from a strong hit from an adult male. I ascribed values of force and speed to individual gestures only after viewing a large portion of the dataset and comparing examples of gestures, ranking each signaller's gestures according to intensity. Those actions that frequently topped the list were ranked as "hard" or "fast." The final variable I used to categorise potential gestures was whether the action could be performed with an object. This variable was important in differentiating certain gestures from one another; i.e. *reach* and *offer* are both visual actions performed away from the body, with medium force and speed, but are distinguished from one another by the fact that *offer* involves the use of an object.

I attempted to use the broadest definitions of different gestures as possible so as to boost the number of examples of each gesture without condensing potentially different gestures into single categories thereby overlooking distinctions that could prove meaningful to the orangutans. To these ends, I avoided making many distinctions between movements based on handshape as handshape appeared to vary wildly depending on signaller, day, and action preceding or following the gesture. Thus, movements such as "touch-with-palm-of-hand" and "touch-with-finger" were both included in "*touch*." The subtler distinction was still encoded under the variable "handshape," but it was not classed as a separate gesture. Conversely, I avoided condensing movements that differed in the quality of movement as I reasoned that much information about the signaller could be extracted by the recipient based on how the signaller moved regardless of whether that information was intended or emotionally-linked. If signallers had learned to use quality of movement to convey different meaning or intent, then we would expect that movements of similar form but different quality would not be interchangeable. Also, if the quality of movement was a direct product of the emotional state of the signaller, recipients might have learned to react to the quality of movement as much as to the type of movement if it was a reliable predictor of how a signaller would act. Either of these scenarios would create a system in which quality of movement was an important variable in determining the reaction of the recipient. For this reason, I used quality of movement as a defining variable for potential gestures.

2.5.3 Determining context of use

Once potential gestures had been defined by their structural elements, the contexts in which they occurred were examined to determine if any were used in exactly the same way. If two gestures that, to us, appear different in some aspect of form occur with the same frequency in the same contexts, it is likely that the difference in form is not a meaningful or salient one to the orangutans. For example, if *grab* and *grasp* are distinguished only by their speed and appear with the same frequency in the same contexts, then speed is not a salient variable distinguishing *grab* from *grasp*. In this example, I would combine the two gestures into a single gesture unless another distinguishing variable could be found.

Contexts were defined by the actions of the most active individual in the dyad. If one individual were resting and the other foraging, the context would be recorded as “foraging.” Context was not considered to be the same as meaning, as gestures that occurred in certain contexts might indicate desires that would not be predicted by the social context. For example, gestures produced during feeding could be requests for food sharing, but could also be requests for the other to move away. Contexts were initially coded in very fine detail, but very few gestures were recorded for some of the contexts and so they were condensed for ease of analysis. For example, the contexts “rest” and “rest-watch” became part of “affiliation,” and “nest-building” was combined with “object manipulation.” The context “affiliation” was later combined with “play” as certain gestures occurred with the same frequency in both contexts, and no salient variable could be found to differentiate on context from the other. The contextual categories included in the end were as follows:

Affiliation/Play – Individuals are socializing in an amicable way or begin socializing within 10 seconds of the gesture. This may involve either tolerance of close proximity during rest or monitoring behaviour, affiliative touching (such as cuddling), or physical play (such as wrestling).

Aggression – One individual lunges at, attacks, or chases another.

Displacement – One individual approaches another and then takes his/her place when he/she moves away.

Fear – Females withdraw from, avoid, or carefully monitor approaching dominant male. This may also be the context for a subordinate following conflict with a dominant or for any member of the group during a human disruption such as a veterinarian visit.

Feeding – One or both individuals ingest food while stationary.

Foraging – One or both individuals search through enclosure substrate (shavings, grass, or bushes) for natural or keeper-supplied foodstuffs.

Grooming – One individual (or both) grooms or is groomed by another.

Intervention – One individual interrupts and stops a dyadic interaction between two others (typically a mother retrieving her infant from another adult or juvenile).

Locomotion – One or both individuals walks or climbs during or within 5 seconds of gesture.

Mating – Individuals in dyad mate during or within 10 seconds of gesture.

Nursing – Infant nurses from mother.

Object Manipulation – One individual (or both) manually or orally investigates, manipulates, or destroys a non-food object (e.g. cardboard box, stick, cloth, bedding or rope). This category includes building nests and making or using tools.

Solicitation – Female displays genitals to dominant male or dominant male investigates female's genitals. Often involves pre- or barely-pubescent females presenting to male.

2.6 Identification of intentional gestures

Video coding of social interactions was focused on identifying physically ineffective movements directed towards other individuals, which could be "potential gestures." In order to further label a particular movement as a "gesture," however, it is necessary to determine whether the individual performing the action is using the movement as an intentionally communicative signal. According to the general criteria previously noted in Chapter 1, an intentional signal must be: 1) directed towards another with 2) the objective of obtaining a goal and 3) be able to be employed flexibly rather than be an automatic response to a stimulus.

In previous studies (Pika et al. 2003; Pika et al. 2005; Liebal et al. 2006; Call and Tomasello 2007), authors claimed that an observation of one individual using a potential gesture in an intentional manner was enough to conclude that that potential gesture was an intentional gesture whenever it appeared within the population. It is, however, an over-interpretation to say that a gesture is intentional for *all* individuals if there is sufficient evidence only for intentional usage in a *few* individuals. Cultural or individual differences may exist in the way apes attribute meaning to movement, as in humans. One woman may lower her eyelids as a meaningful invitation to a potential suitor whilst another may never use the movement purposefully, lowering her lids only when tired. Since each individual may differ in the degree to which they assign meaning to a particular movement, it is necessary to demonstrate that each individual uses each movement as a gesture. If, however, we can identify at least one example of a gesture used by an individual that strongly suggests intentional use, we may assume that other uses of that movement by that individual are also intentional. If evidence for intentional usage were required for each "utterance," then most human communication would likely be discarded due to lack of evidence.

2.6.1 Determining intent and meaning: new variables

In order to determine which potential gestures (already identified as a social and not physically-effective movement) were true gestures, the video clips were re-examined and subjected to a second round of coding. Five further variables were coded in order to help identify cases of likely intentionality. These additional variables were:

1) The presumed goal of the signaller.
2) The type of response from the recipient.
3) Whether or not the response appeared to fulfil the goal.
4) The eventual outcome of the exchange from the perspective of the signaller.
5) A rating of how well the clip satisfied the criteria for intentional usage.

The possible values for each of the variables are enumerated in Table 4. The five new variables correspond to: 1) "goal," 2) "response," 3) "goal met?," 4) "outcome," and 5) "intentionality rating," in Table 4. Table 4 contains more than the five new variables as it covers all variables used to determine inter-observer reliability.

The presumed goal of the signaller was a subjective measure based on the context and the frequency of certain outcomes ending the signaller's persistence. If an individual often produced a gesture in a particular context and ceased further communication when the recipient responds in a certain way, I assumed that the signaller's goal was to cause that response in the recipient. This logic does not hold for types of gestures that are often met with agonistic reactions such as the recipient's refusal to share an item or the recipient forcibly rejecting the signaller. Also, since I assumed that the goal of a gesture is to effect a change in the recipient, gestures that frequently failed to elicit a reaction from the recipient were not considered to be without a goal. Many infant gestures made during feeding or resting contexts are ignored or actively rejected by adults. If the presumed goal of the signaller were determined solely by the recipient's most frequent reaction, many infant gestures would be deemed to be requests for the other to reject or ignore the infant.

Response was a measure of the overall tone of the recipient's response to each gesture. The specific action performed by the recipient (e.g. "push") was categorised based on its affiliative or agonistic nature. Thus "push" would likely be recorded as a negative response. Response categories aimed to determine whether the recipient responded negatively, positively, paid attention to the signaller, or did not respond at all. Also recorded was whether or not the type of response appeared to fulfil the goal of the signaller. If the response appeared to fulfil the goal, the expectation was that the signaller would stop signalling, whereas if the response did not fulfil the goal, the signaller would likely continue trying to reach her goal.

The "outcome" is distinguished from the response as it is the eventual result of the entire interaction from each individual's standpoint rather than the recipient's immediate reaction to each gesture. Outcome is intended to measure the net result of the interaction, so it may be positive for one individual and negative for the other.

As an illustration of the application of the four variables discussed above, consider the hypothetical following encounter: individual A approaches individual B with a food-begging gesture and individual B waves the back of her hand toward individual A, causing him to pause slightly. Then individual A grabs the food B is holding and B lunges at A with her teeth bared. A then leaves. In this example, the *goal* for the food-begging gesture and the grab produced by A would be "food share," the *response* would be "negative" and would not fulfil the goal, and the *outcome* would be "reject." Taken from the standpoint of individual B, the *goal* of both the hand wave and the lunge would be "stop," the *response* would be "none" for the hand wave and "negative" for the lunge; the response to the lunge would fulfil the goal, and the *outcome* would be coded as "leave." In this type of coding, a communicative interaction can be successful for one individual and unsuccessful for the other.

2.6.2 Rating gestures for intentionality

An “intentionality rating” was given to every instance (token) of each potential gesture based on how strongly it suggested that the signaller had both an intended recipient and specific goal. I used the gaze of the signaller prior to gesturing to determine whether visual and auditory gestures had an intended recipient. Tactile gestures were directed at one specific individual by definition. In order to establish whether the signaller had an intended goal, I looked for evidence that the signaller expected a response from the recipient. Measures of expected response included response waiting, gaze alternation, persisting, and using modalities appropriate to the attentional state of the recipient (e.g. visual gestures when the recipient is looking).

Based on the above measures of signaller intent, I rated each gesture according to how well it demonstrated intentional usage. Every example of a potential gesture was given a rating on a 4 point scale corresponding to values of: 1) no evidence for intentional use (e.g. performing a potential gesture away from another individual), 2) ambiguous use (e.g. looking toward the recipient but showing no sign of expecting a response), 3) consistent with intentional interpretation, and 4) support for intentional interpretation (i.e. gesture clearly directed towards another with strong evidence that the signaller expects a response from recipient). Each gesture clip was given an intentionality rating on a scale of 1-4 according to the following guidelines:

1. The gesture appeared to be non-communicatory (e.g. it was performed without any other individual present), or if it was an effective action rather than a gesture.
2. The gesture occurred in the presence of another individual, but had no other signs of communicative or intentional use.
3. The gesture occurred in the presence of and was directed towards another individual.
4. The gesture was directed towards another individual and the signaller indicated that a response was expected by waiting for a response, looking towards the recipient's face, escalating the signal, or following the gesture with either another gesture or an effective action.

In compiling gestural repertoires for each individual, I included only gestures that that had been observed at least once in clearly intentional usage (a ranking of 4) or which were used at least 50% of the time in a manner consistent with an intentional interpretation (a ranking of 3). Potential gestures were discarded from an individual's repertoire if there were no examples of use earning a ranking of 3 or 4. If there were no examples of intentional use by any individual at one zoo, that gesture was discarded from the set of gestures observed in that group. Comparison of gestures used by individuals of different ages and at each of the 3 zoos is included in Chapter 3. A table showing all intentional gestures used by each individual is provided in Appendix IV.

2.7 Inter-observer reliability

2.7.1 Secondary observer

Inter-observer reliability for identification of gesture form, content, and intentionality was measured by having a second rater code a subset of the video clips using a modified coding spreadsheet. This rater, Cat Hobaiter, was familiar with the gestural communication of gorillas and had coded many videos of captive gorilla interactions. She therefore provided an ideal second rater as she was used to the theory and methodology but was unfamiliar with the species used in the study and thus had no preconceptions of orangutan communication.

2.7.2 Design

The coding sheet used for reliability tests was modified from that used for the original coding of gestures to accommodate the rater's unfamiliarity with both the individual orangutans included in the study and the behaviour of the species. The reliability coding aimed to focus on the elements of the videotaped actions essential to determining their status as gestures as well as their supposed meaning. The coding of the second rater, therefore, consisted primarily of variables used to determine intentionality, the supposed goal of the action, and the response of the recipient. A full list of variables coded is given in Table 4 and an example of an entry in the spreadsheet is included in Appendix III.

Table 4: LIST OF CODED VARIABLES USED IN ANALYSIS OF INTER-OBSERVER RELIABILITY.

Variable	Type	Possible values
Mechanical effectiveness	Scale 1-4	1) Effective 2) Possibly effective 3) Likely non-effective 4) Definitely non-effective
Directedness	Scale 1-4	1) No recipient 2) Several potential recipients 3) Several potential recipients but directed to one 4) One definite recipient
Goal	Categorical	• Unknown • Affiliation • Attention • Play • Share food/object (acquire object or info) • Look at object/body part (direct attention) • Stop behaviour ("no") • Move back • Leave • Follow • Climb on • Pick up • Mate
Signaller's visual attention	Out of view & Scale 1-4	• Out of view 1) Can't see recipient 2) Can potentially see 3) Looking towards 4) Looking at the face or eyes
Recipient's visual attention	Out of view & Scale 1-4	• Out of view 1) Can't see signaller 2) Can potentially see 3) Looking towards 4) Looking at the face
Modality match	Categorical	• Not detectable • Detectable but not necessary • Detectable and necessary
Response waiting	Scale 1-4	1) None 2) Pause 3) Wait until response 4) Wait >2 sec
Response	Categorical	• No response • Negative (look away, move away, aggression) • Acknowledge but carry on with prior behaviour • Pay attention (look or move towards) • Positive interaction (affiliate, play, give)
Goal met?	Categorical	• No • Yes • Unclear
Persistence	Categorical	• None • Repeat/elaborate • Same modality • Change modality

Table 4 continued

Variable	Type	Possible values
Sequence goal	Categorical	- Different - Same - Unclear
Outcome	Categorical	- None - Affiliation - Attention - Play - Share food or object - Look at object or body part - Stop behaviour - Move back - Leave - Follow - Climb on - Pick up - Mate
Intentionality rating	Scale 1-4	1) Not intentional 2) Unclear/needs more evidence 3) Consistent with intentional interpretation 4) Support for intentional interpretation

2.7.3 Procedure

The second rater (CH) was trained to use the spreadsheet on a set of 15 pre-selected video clips. The primary rater (EC) analysed the 15 clips alongside the second rater, discussing why each judgment was made and working with one clip until both agreed on all the different ratings. Then the second rater was given free access to all video clips and told to code as many clips as possible within a limited period of time (two afternoon sessions). She was given no other instructions or limits except that she should include some video clips from each of the three zoos.

2.7.4 Analysis

Tests for reliability between the observations of the two observers were done using Cohen's Kappa. This test measures the agreement between independent observers, taking into account the possibility of chance agreement.

Some of the variables were combined into more general categories to reflect the overall nature of the interactions rather than highly specific distinctions between contexts that may require familiarity with either orangutan behaviour or the ability to contextualise the subset of clips within all clips in the dataset. Thus reactions that involved non-aggressive social interactions were grouped together as “positive” responses, and responses that involved leaving or actively rejecting the signaller (e.g. pushing away) were combined into “negative” responses. I grouped actions that were coded as scalar values into 2 categories of high and low values for analysis. The combining of specific values for each variable is reported below.

Values for the variable *mechanical effectiveness* were combined into either “effective” (previously, “effective” and “probably effective”) or “non-effective” (previously “likely non-effective” and “definitely non-effective”). I condensed the category *directedness* by combining “one potential recipient” and “one certain recipient” into “one recipient.”

The variable *goal* was condensed into more general categories that reflected either attraction or repulsion of the recipient. The values “leave,” “stop,” and “move back” were combined into “stop/move away.” The values that reflected the goal of positive interaction (“affiliation,” “attention,” “play,” and “look at body part”) were combined into “attention/play.”

The measures of gaze direction for both the signaller and recipient (*signaller visual attention* and *recipient visual attention*) were collapsed within each variable so that all values that indicated one individual could see the other became “looking towards.” Thus measures of visual attention had values of either “looking” or “not looking.”

Response waiting was initially divided up into 4 categories in order to obtain a more delicate measure of whether the signaller was demonstrating her expectation of a response from the recipient. For the purposes of this analysis, only the most extreme measure of waiting for a response (waiting for more than 2 seconds) was counted as response waiting. All values indicative of pauses shorter than 2 seconds were condensed into "no response waiting."

For analysis of the variable *response*, the value "acknowledge but carry on with prior behaviour" was merged into the value "pay attention to." The variable *outcome* was condensed using the same combination of categories as was used for *goal*.

The rating for *intentionality* was condensed so that both of the values that suggested intentionality ("consistent with intentional interpretation" and "support for intentional interpretation") were merged into the single value "likely intentional." This was done to reflect the inclusion of both values in building the dataset of intentional gestures.

2.7.5 Results

The second rater coded 64 video clips, yielding a total of 108 potential gestures. Nineteen of the potential gestures had to be discarded due to incomplete coding. This left 89 potential gestures (5.8% of all gestures) to use for comparison of the two raters. The kappa values for concordance between the two raters are reported in Table 5.

Table 5: MEASURES OF AGREEMENT (COHEN'S KAPPA) BETWEEN THE TWO OBSERVERS FOR EACH OF THE 13 VARIABLES MEASURED.

Also listed is the type (scalar or categorical) for each variable. The strength of agreement signified by each kappa value (Landis and Koch 1977) is given in the right hand column.

Variable	Type of variable	Kappa value	Strength of agreement
Mechanical effectiveness	Scalar	.88	Almost perfect
Directedness	Scalar	.94	Almost perfect
Goal	Categorical	.63	Substantial
Signaller's visual attention	Scalar	.91	Almost perfect
Recipient's visual attention	Scalar	.89	Almost perfect
Modality match	Categorical	.78	Substantial
Response waiting	Scalar	.79	Substantial
Response	Categorical	.64	Substantial
Goal met?	Categorical	.48	Moderate
Persistence	Categorical	.80	Substantial
Sequence goal	Categorical	.83	Almost perfect
Outcome	Categorical	.56	Moderate
Intentionality rating	Scalar	.68	Substantial

Though the values for two variables generated only "moderate levels of agreement," the mean kappa value for all variables was 0.75, signifying a "substantial" strength of agreement between the two raters.

Chapter 3: Gesture form and function

3.1 Introduction

All types of movement can carry meaning. Involuntary reflexes such as crying or smiling transmit information about the emotional state of the individual who displays them. Functional actions such as walking or scratching can be used to determine the mood of an individual or to predict their subsequent actions. Both involuntary and functional actions may function as meaningful units of movement, but should not be considered gestures. For a movement to be included as a gesture in this study, it must be performed as a communicative act. I aimed to define and investigate gesture from the signaller's perspective, focusing on the form, function, and use of gestures. I believe that gestures must be defined from the signaller's perspective—it is only through the signaller's intention that a movement becomes a meaningful gesture. This distinction allows us to examine a relatively small subset of movements that can be used to gain information about the signaller or the environment. This narrowing down of interest focuses our investigation on movements that may be conventionalized signals: actions that take on meaning through their use as signals rather than by any intrinsic property of the actions themselves. These signals have the potential to be both produced and understood by any individual.

3.1.1 Previous studies

Relatively few studies of primate communication have focused on orangutans, and relatively few studies of orangutans have focused on communication. There are only two previous studies that focus on the gestures of orangutans. In the rest, gestures are mentioned as part of a larger body of social behaviour.

Wild populations

Previous studies of orangutan behaviour in the wild have mainly catalogued communicative gestures and body postures as a subset of all behaviour (MacKinnon 1974; Rijksen 1978). These studies aimed to catalogue and explain general orangutan behaviour rather than focusing on the structure and use of gesture specifically. The predominantly solitary nature of wild orangutans does not lend itself easily to the study of social communication. The arboreal nature of the animals makes it difficult to observe them from close distances, making study of visual signals even more difficult. The only social pairing that can be reliably found in wild orangutans is the mother-infant dyad. It is not surprising then that the only study of free-ranging orangutans to focus on the structure and use of gesture investigated infants' gestures toward their mothers (Bard 1992). This study focused on the development of infants' intentional requests for food and on the ability of the infant to manipulate a social agent in order to obtain a specific goal. Bard found that infants progressed from functional actions directed towards food items (such as grasping and pulling) to non-functional gestures directed towards the mother (such as reaching or touching the mother's hand) as they aged. This finding supports the theory that gestures are ontogenetically ritualized from effective actions (see Call and Tomasello 2007), but the study was restricted to a highly specific context and may not be representative of the development of other gesture types.

Captive populations

There has only been one previous study that has focused solely on the gestures of captive orangutans. This study, led by Katja Liebal (2006; 2007), was part of a comprehensive project by the Max Planck Institute for Evolutionary Anthropology to compare the presence and use of gestures between all the apes (for an overview of the entire project see Call and Tomasello 2007). These studies catalogued the gestures seen in captive populations of each kind of ape and analysed use of gestures based on recipient's visual attention, age and sex of recipient, and group membership. By using identical criteria across species to determine which movements should be classified as gestures, this set of studies was able to directly compare differences in gestural repertoires between species. While they found a few differences in the proportion of visual and tactile gesture between species, they did not observe any significant differences in overall species repertoire size or use.

In her study of two orangutan groups, Liebal (2006; 2007) identified 14 tactile gestures, 15 visual gesture, 5 facial expressions, and 10 "actions." For a type of movement to be counted as a gesture in her study, examples of the gesture had to be both directed towards other individuals and used flexibly in relation to context. Specific criteria for inclusion as intentional gestures were:

> *Intentional communication was defined as a motoric act directed to a recipient via body orientation, eye gaze or physical contact with the sender expecting a response as evidenced by looking to the recipient, waiting for a response or persisting in the communicative interaction (Tomasello et al. 1985; Tomasello et al. 1994; Sarimski 2002). Applying the following three criteria, an observed behavior was thus defined as an intentional signal if it (1) was observed at least two times over the whole observation period (which ensured that this gesture served to reach a recurrent social goal); (2) was directed at a particular recipient; and (3) was used flexibly in different social contexts, or else several signals were all used in the same context ['means-end dissociation' (Bruner 1981)].*

While this definition aims to ensure that movements that are counted as gestures are 1) frequent, 2) directed, and 3) flexible, data from all individuals were pooled together. Thus, a movement was considered to be a gesture if any individual used it according to their definition of intentionality. This approach treats all individuals as one, and assumes that if a gesture is used intentionally by one individual, the same movement in other individuals must be an intentional gesture as well. Their criteria for intentionality aimed to restrict gestures to those that were actively chosen by the signaller to communicate with another, but their application of these criteria did not allow for differences between signallers. This presents a serious shortcoming in their interpretation of the use of gestures in different populations and in different contexts. It is possible that one individual figured out how to use a movement in its behavioural repertoire to elicit a particular response from others, but that other individuals perform the same behaviour without the attachment of meaning and expectation. It is also possible that if one observes a behaviour used intentionally by one individual, then all individuals that express that behaviour must be doing so intentionally as well. Since both of these possibilities exist, it is appropriate to take a more rigorous approach to determining which movements should be counted as gestures. The approach I employed of ensuring that each signaller had to exhibit at least one highly intentional usage of a particular gesture type before it was included in its repertoire addressed this concern.

All of the studies of gesture performed by researchers at the Max Planck Institute in Leipzig analysed the use of gestures in different contexts. Liebal et al. (2006; 2007) described the use of gestures in different "functional contexts" including "Access, Affiliation, Agonism, Grooming, Ingestion, Parental care, Play, Sexual behaviour," and "Unknown." However, the authors use contextual occurrence as an equivalent measure for the signaller's goal. They list the contexts in which each gesture was seen to occur and then use multi-contextual use as evidence of intentional usage. This approach is useful in determining how many types of gestures are used during different social contexts, but does not address the question of meaning. Gestures that occur in the same context do not necessarily have the same goal. In the context of play, some gestures may be used to initiate play, some to refuse play requests, and some to end play in progress. All of these gestures are used in the play context, but their goals are more important than the context. Gestures used to initiate play may be used in the context of play, but they may also be used to initiate play during foraging or to elicit affiliation during an attempted displacement. Gestures used to refuse play may be used to refuse requests to be carried or to share food as well. Contextual use is illustrative of the goals that occur during different contexts, but does not address signaller intent during gestures. The authors of the studies on ape gesture originating from the Max Planck in Leipzig used "the disassociation of means and ends" as a measure of the intentionality of gestures (see Call and Tomasello 2007). Disassociation here refers to a single gesture type being used in several different contexts or several different gesture types being used in the same context. Context was used as a measure of the goal or “end” of gestures. Since goals may differ strikingly from context, gestures that appear in several contexts may, in fact, have the same goal. An analysis of means-ends disassociation using context as a proxy for goal may therefore lead to spurious conclusions about the flexibility of certain gestures.

3.1.2 Let the gesture do the talking

In designing this study of orangutan gestures, I wanted to begin with all social non-functional movements and let the similarities in their structural elements and the differences in their use determine which movements should be grouped together as gestures. This bottom-up approach to assembling a repertoire of meaningful, intentional gestures highlights the signaller’s intentions and the communicative process rather than merely cataloguing the actions observed. I used structural elements to assemble likely gestures (similar to previous studies), but then tested my set of gestures by examining their ability to predict observable goals, a measure incorporating both the goal of the signaller and the outcome of the interaction. This approach transformed what was previously a fairly subjective process of defining gestures into a more objective one, defining gestures by prioritizing different types of structural similarity and using observed goals to validate gestural definitions. By incorporating both structural and motivational elements into the definitions of gestures, it is possible to address questions about the interaction between communication and cognition.

Though the bottom-up approach did not always yield well-defined gestures with one-to-one meanings, it was able to identify many meaningful gestures, and provide insight into which variables are most important to orangutans when choosing how to communicate certain desires. The strength of my process of identifying gestures lies not in the ability of the observer to choose which gestures appear communicative, but in allowing the use of movements by individual animals to indicate what movements are important in their communication system, and what variables are important in creating those movements. This approach, though initially a lengthy process, allows the gestures themselves to "do the talking" and removes much of the subjectivity from the identification and interpretation of gestures. The analyses presented in this chapter are not limited to the study of orangutans or apes. The strength of the approach I used is that it allows one to investigate the communication system of any group or species with which one shares no language by focusing on the use of the system itself and building a set of communicative units that are likely to be meaningful to the individuals observed.

3.1.3 Methods

In order to identify movements that potentially held meaning as gestures for the signalling orangutans, I began by selecting all non-functional movements I observed orangutans perform when oriented towards other individuals and classifying them according to their forms. The movements were categorized according to their modality, direction, speed, force, and whether or not they were performed with an object (according to the classification in Table 3). These distinctions allowed movements to be sorted into potential gestures through similarity of form.

Similarity of form does not necessarily imply similarity of meaning, as English homophones and homonyms demonstrate. In order to demonstrate that two potential gestures that were similar in form were distinct signals, it was necessary to ascertain that 1) the same individuals could use both gestures (ensuring one gesture was not a reflection of the movement preferences of certain individuals), and 2) that the two gestures were not used interchangeably towards the same goal. If one of the two potential gestures was always used first and the other always followed, or if one were used to initiate a particular kind of interaction with an adult and the other with a juvenile, then it would be possible to argue that the two potential gestures were indeed different from one another. In these cases, however, one would have to be careful that one of the two forms was not an exaggeration of the other, following the other as a necessity because it was used when the first gesture failed to achieve its goal and was repeated in a bigger or slower way. If this second form could precede the other or be used independently, it would indicate that it was a distinct gesture, but I tried to ensure that gesture types were defined more often as categorical distinctions than gradations of the same movement along one particular variable.

In order to investigate the choices orangutan signallers make and what variables they take into account when attempting to communicate, I first had to establish that the potential gestures they produced towards other orangutans were, in fact, intentionally communicative signals. As I planned to investigate orangutans' use of gesture as a way to gain insight into their understanding of communicative failure and the minds of other individuals, it was important that the signallers actively chose the gestures they exhibited. Intentional use was therefore essential in determining which movement forms (i.e. potential gestures) could be considered to be gestures in this study. To this end, I identified likely markers of intentional communication (i.e. waiting for a response, looking at the recipient following a potential gesture, using a modality that can be perceived by the recipient, persisting and elaborating when there is no response) and gave each occurrence of each potential gesture a rating to indicate how strongly it was accompanied by these markers.

In the analysis of the forms and meaning of gestures presented in this chapter, I focused on gestures that were produced in an intentional manner at least once by each individual observed to use that gesture (see Chapter 2.5 for a full description of intentionality criteria). It is possible that though a particular movement may be the same in two individuals, one may use the movement with an intent to communicate and one may not. It is for this reason that before I claimed that an individual had a specific gesture in her repertoire, she must have used that gesture at least once in a manner that suggested intentional use.

Non-functional movements that were performed towards other individuals (or at least in the presence of other individuals) were initially divided into potential gestures based on the modality and overall form of the movement (Table 2). Once I had identified the gestures that were used in a manner suggestive of intentional communication, I then examined the contextual use, presumed goal of the signaller, and response of the recipient of each example in order to determine whether the gestures that I had defined through structural similarity were also distinct from one another in their contextual use as signals.

The presumed goal of a gesture was taken to be the action of the recipient that caused the signaller to stop trying to communicate. I assumed that the goal was always to elicit an active response from the recipient (e.g. movement towards or away from the signaller, sharing an item, or interacting in an affiliative way). Thus, a lack of response of the recipient, or a rejection of the signaller (by averting gaze or turning away) was never assumed to be the signaller's goal. These presumed goals were compared within and between gestures in order to best categorize them into potential gestures. Some gestures, which had appeared distinct in form, were combined into larger units because they were used interchangeably. Others were divided into more than one gesture if the variance in their use could be explained by separating out a subgroup defined by a commonality along a structural variable such as handshape or location on the recipient's body. These combinations and divisions helped to ensure that the observed movements grouped into potential gestures were similar in form but were used to convey as few goals as possible.

If gestures were used interchangeably towards any goal, then gestures could not be said to carry any specific information about the signaller's goal. If this were the case, all the gestures would signal is intent to communicate rather than providing any specific information about the goal of the signaller. Following one of these gestures, the recipient would have to discern the signaller's goal from non-gestural cues such as context. This is the case in the deictic[1] gestures (such as pointing) produced by adult humans and children (McNeill 1992). When a young child points to an object or individual, the recipient must interpret the child's specific meaning from the context. Apes are capable of similarly ambiguous attention-directing gestures, though the nature and form of "pointing"

---

[1] A deictic signal is one in which meaning is dependent on the context. Pointing is perhaps the most common instance of deixis; the movement itself does not transmit information, but through using the signal in conjunction with the context, an observer can infer meaning. English deictics exhibit similar traits, including words such as "there" and "that."

gestures in apes is debated (see Leavens and Hopkins 1999 for review). Other attention-directing signals may be more common in ape gestural communication. One study of human-reared gorillas found that they would use “contact gestures” (e.g. directing the human’s hand to a location of desired food) in situations where human infants would have used pointing to indicate intent (Gómez et al. 1993)

In contrast to deictic gestures that rely on the context for meaning, gestures that are used most frequently to initiate one type of interaction likely contain information about the signaller’s specific goals in the form of the gesture. I attempted to define gestures so that they corresponded to as few goals as possible. This was done to raise the specificity of each gesture so that meaningful gestures could be identified whilst ensuring that the definitions were not so narrow as to render most gestures idiosyncratic.

3.2 Identifying intentional gestures

Using the criteria for determining intentionality explained in section 2.5, I narrowed down the species repertoire to those gestures that had been observed in a manner highly suggestive of intentional use by at least one individual. I insured that each individual who was said to have a certain intentional gesture in her repertoire had demonstrated intentional use of the gesture at least once. If an individual performed a potential gesture that was used by others in an intentional way but demonstrated no intentional usage herself, all examples of that individual using that movement were excluded from the dataset.

These criteria pared down the initial dataset of 1581 potential gestures to 1334 intentional gestures. The distribution of intentional gesture use by all individuals in the study is given in Appendix IV. The exclusion of non-intentional potential gestures caused nine potential gestures to be deleted from the dataset, leaving 62 gestures that met the criteria for intentional usage. The potential gestures *chin-up/nod, clap, investigate, long-body-scratch, raspberry, rock, smile-with-teeth, top-lip-out*, and *wipe-face* were all removed from the dataset because they were never observed to be used intentionally towards conspecifics.

The two auditory gestures, *clap* and *raspberry* were used extensively with keepers and were employed frequently during the experiment described in Chapter 5, but they were observed only rarely during interaction with other orangutans and there was no evidence that they were used intentionally toward conspecifics. These two gestures were therefore excluded from the analysis of natural orangutan gestures, but were included as gestures in the experiment presented in Chapter 5 as they were directed towards the experimenters and met other criteria for intentionality when produced under the experimental conditions. Following the removal of *clap* and *raspberry* from the dataset, the other auditory gesture, *hit ground/object*, was added to the set of visual gestures so that it could be included in further analysis based on modality. As this gesture had a highly salient visual element, I felt that it functioned as both a visual and auditory gesture.

One of the potential gestures, *investigate*, was deemed to be effective in achieving its goal (to inspect a body part of another individual) rather than communicating a goal to the recipient and so was discounted as a gesture. The two facial expressions (*smile-with-teeth* and *top-lip-out*) in the list were each observed in only a single individual, and while they were directed towards others, were never accompanied by response waiting, elaboration, or other measures of intentionality. *Long-body-scratch* and *rock* were observed in only one or two individuals each and appeared to be undirected actions exhibited prior to or following agonistic encounters. It seems likely that both were used as repetitive stress-relieving actions rather than communicative signals.

The last gesture excluded, *wipe-face*, was observed in only one individual and seemed to be associated with stress as well. The female that used this action performed it often when watching the dominant male, but did not direct other gestures towards him during these encounters or display any measures of response expectation, so it was deemed to be an undirected action and likely a sign of nervousness. Both the form and context of *wipe-face* are similar to the “muzzle-wipe” gesture found in baboons (Wallis 2004). It is possible that the similarity of these gestures is indicative of both a common emotional state and muscular reflex to that state. But as *wipe-face* was only observed in a single orangutan and is relatively common in baboons, it seems unlikely that the gesture is shared by the two genera. The observed similarity of context does, however, imply a similar emotional state in the two species, and may be an indication that the orangutan who exhibited the gesture was experiencing unusual stress.

3.3 Sensitivity to gaze

One of the most important social variables in determining the success of a gesture is whether or not the recipient can perceive it in the first place. Visual gestures performed when or where the recipient cannot see the signaller have no chance of being successful, as they are not received. This distinction moves beyond simple audience effects, taking into account the perceptual abilities of the recipient and the modality of the communicative signal. Previous studies have shown apes to be more likely to use communicative signals when or where they could be potentially received (Liebal et al. 2004; Poss et al. 2006; Call and Tomasello 2007). I analysed the use of visual and tactile gestures to recipients who could or could not see the signallers and found that my data supported these previous findings.

In the 3 populations I studied, tactile gestures were used more frequently than visual gestures in general (62% of the time). Gestures of both types were produced more frequently to recipients who were attending than to those looking away (64% of the time). However, 75% of visual gestures were used towards attending recipients, whereas only 58% of tactile gestures were. The number of tactile and visual gestures we observed in both conditions is illustrated in Figure 2. Using a chi-squared analysis, I found a significant difference between the use of visual and tactile gestures based on the recipient’s visual attention ($\chi^2 = 41.7$, df = 1, $p < 0.01$). I also performed binomial tests on the observed frequencies of visual and tactile gestures directed towards recipients who were looking based on the average probability (.64) of being performed towards a looking recipient. Visual gestures were used more frequently than chance towards recipients who were looking (Binomial test, n = 521, p < 0.001). Tactile gestures were used at chance level. This indicates that signallers are sensitive to the gaze of their recipients and are choosing gestures that can be perceived by the recipient. This finding agrees with the finding of Liebal et al. (2006) that orangutans perform visual gestures more frequently when the recipient can see the signaller.

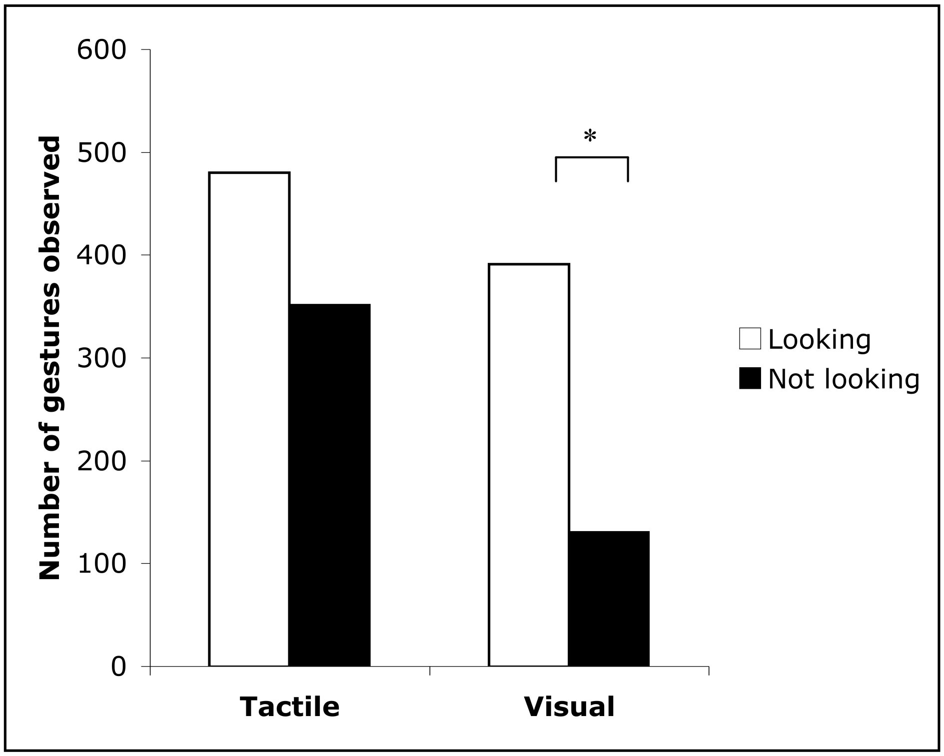


Figure 2: GESTURAL MODALITIES USED TOWARDS RECIPIENTS WHO ARE LOOKING OR NOT LOOKING TOWARDS THE SIGNALLER.

Visual gestures were used significantly more often when the recipient was looking. The dataset of 1421 examples of intentional gestures (and all gestures in sequences) was used for this analysis.

3.4 Ascribing meaning

To determine whether the observed gestures reliably carried any meanings, it was necessary to compare the goals of each use of a particular gesture type to determine whether the gesture predicted a specific response. Towards this end, I combined two measures of meaning in order to achieve a variable of meaning that was objectively based on the recipient's reaction to the gesture, but retained some measure of the signaller's intent. This was done by analysing meaning only for gestures whose supposed goals (a fairly subjective measure) matched their outcome (an objective measure based on the outcome of the interaction). These gestures could be considered to be those that appeared "successful," in that their goals were achieved. By using only successful gestures in the analysis, I avoided the problem that many orangutan gestures received a negative or repulsive reaction in which the recipient pushed the signaller away or turned away.

In the analysis of meaning, I assumed that the goal of gestures was always to effect a change in the state of the recipient and that looking away or physically rejecting the signaller was never the goal. By excluding these types of reactions from the possible set of goals and restricting the database to those gestures whose outcomes matched their goals, I ensured that all examples used to determine meaning were aimed at eliciting (and achieved) active responses from the recipient. The changes in the recipient (i.e. eventual outcomes) that I considered possible goals included affiliative, agonistic, sexual, and acquisitional desires (Table 6).

Table 6: Presumed goals of communicative interactions.

Type of Goal	Goal
Attraction	Affiliation/Play
	Attention (visual)
Directing attention	Look at object
Obtain item	Share food or object
Coordinated Locomotion	Climb on
	Pick up
	Follow
Repulsion	Move back/Leave
	Stop (current behaviour)
Solicitation	Sexual behaviour

3.4.1 Presumed goals

The presumed goal was coded for each occurrence of each gesture based on context, recipient's behaviour prior to the potential gesture, and final recipient response (where the signaller did not persist). Presumed goals were assumed to be either active interactions (play, pay attention, pick up, mate), acquisition of items (share food or object), movement towards (follow), or movement away (leave). Though this measure was reliant on some degree of experimenter subjectivity, an inter-observer reliability comparison of two experimenters' attributions of signallers' goals yielded a kappa value of .83 (see Chapter 2.7), signifying a high level of agreement. Therefore, presumed goal was concluded to be a replicable measure and useful tool in further analysis of gesture meaning.

3.4.2 Observable goals

An occurrence of a gesture was deemed to have an observable goal when the eventual response from the recipient was the same as the signaller's presumed goal (i.e. if the gesture or sequence was "successful"). While the presumed goal was a useful tool to help ascribe intent to the signaller in each use of a gesture, its measure was somewhat subjective. To discuss meaning of particular gestures, it was necessary to adopt a more objective stance. By reducing the dataset to cases where the presumed goal was the same as the outcome of the interaction, the presumed goals were supported by the recipient's response. Since observable goals require a match between goal and outcome rather than being solely based on recipient response, no gesture was ever assumed to have the goal of producing no response or of averting the attention of the recipient. The eventual outcome of the interaction rather than the immediate response of the recipient was used as the objective measure of recipient response to account for sequences of gestures and "dialogues," which occurred frequently. A sequence was defined as one individual's gestures directed towards another that occurred without a break of more than 10 seconds between gestures (sequences are explained further in Chapter 4). The goals of all gestures in a sequence were not necessarily assumed to be the same, but the outcome of a sequence was shared by all gestures in the sequence. I designed the criteria for observable goals to account for multiple gestures in a sequence as well as the outcome from each participant's perspective. Consider the following sequence between orangutans A and B:

A: ***NUDGE*** $_{\text{gesture}}$
B: no response $_{\text{action}}$
A: ***GRAB*** $_{\text{gesture}}$
B: ***TOUCH*** $_{\text{gesture}}$
A: ***PUSH*** $_{\text{gesture}}$
B: leave $_{\text{action}}$
Result: B leaves

In this example, the presumed goal for Orangutan$_{\mathbf{A}}$'s gestures would likely be "move back/leave" and the outcome would be "move back/leave." The presumed goal for Orangutan$_{\mathbf{B}}$'s gesture would likely be "stop" and the outcome would be "none." Thus, Orangutan$_{\mathbf{A}}$'s gestures "*nudge*" "*grab*" and "*push*" would have the observable goal "move back/leave" and Orangutan$_{\mathbf{B}}$'s gesture "*touch*" would likely have the presumed goal "stop" but have no observable goal, as the outcome did not meet the presumed goal.

By restricting analysis of meaning to gestures with observable goals, I reduced the dataset from 1334 intentional gestural tokens to 698 intentional tokens with observable goals. Of the 646 gesture tokens whose outcomes did not match their goals, 344 received no reaction as the outcome, and 99 ended in rejection (i.e. the recipient turned away or pushed the signaller away). This meant that only 203 tokens ended in an active outcome that did not match their supposed goal. These 203 gestural tokens with goal-outcome mismatches could have been caused by misinterpretation of the presumed goal by the experimenter, or could have been an effect of a communication system in which recipients do not always respond "appropriately." The relatively low frequency of "mismatched outcome" gestures to "observable goal" gestures led me to accept the use of observable outcome as a proxy for "meaning" in the analysis of semantic content of specific gestures, as it included most of the gestures that received an active outcome.

For all gestural tokens with observable goals, the outcome of the interaction matched the presumed goal except when the presumed goal was to gain the attention of the recipient. For interactions where the presumed goal was to attract attention, all affiliative outcomes that satisfied the presumed goal were counted as matching even if they included other actions, since the recipient almost universally directed their attention towards the signaller before engaging in an affiliative interaction. Thus, affiliation/play and look at object/body part were counted as satisfying the presumed goal of attracting attention.

Gesture types that were found to have at least one frequent goal were further analysed to determine whether or not they were used to predictably elicit a specific response from the recipient. The reduction of the dataset to gestures with observable goals initially left us with 62 intentional gestures. However, 24 of these gestures were used fewer than 4 times towards any one goal and were thus not included in further analysis as they were deemed to be too infrequent, and were not observed or were not successful frequently enough to allow meaningful comparisons of the signallers' observable goals. Once the 24 gestures were excluded, 38 gestures remained that were used frequently in conjunction with at least one observable goal. These 38 gestures and their contextual use are given in Table 7.

Table 7: USE OF GESTURES WITH OBSERVABLE GOALS TOWARDS GOALS OF DIFFERENT TYPES.

The table includes only those 38 gestures that were found to have an observable goal at least 4 times in a single context. The 38 gestures were used a total of 639 times in conjunction with an observable goal. Numbers in the cells are the number of observed cases that a particular gesture has been used successfully (i.e. met the criteria for an observed goal) towards that goal.

Gesture	**Affiliate/play**	**Attention**	**Climb on**	**Follow**	**Look at object/ body part**	**Move back/ leave**	**Pick up**	**Share food/object**	**Stop**	**Total**	**% Towards one goal**	**Ratio primary: secondary goals**
grasp	15*	2	5	3		2	3	5	8	43	35%	1.9
pull	18	5	6	29*		1		16	3	78	37%	1.6
turn head	5*	4						4		13	38%	1.3
grab	22*	2	1	1	1	4		7	10	48	46%	2.2
mouth	3					1		4*		8	50%	1.3
touch	28*	3	1	5		7		6	3	53	53%	4
brush	1	1		1		2			7*	12	58%	3.5
swat	13*	1				6			2	22	59%	2.2
pull away	1					3			6*	10	60%	2
reach	16*		1	2	1		3	2		25	64%	5.3
push	6					30*		1	7	44	68%	4.3
air bite	5					1			1	7	71%	5
bite	13					2			3	18	72%	4.3
embrace	8		1	1		1				11	73%	8
nudge	2					19			5	26	73%	3.8
hit	28	1				2			6	37	76%	4.7
back roll	4	1								5	80%	4
shake	4	1								5	80%	4
offer	1				5					6	83%	5
pull hair	11					1			1	13	85%	11
embrace pull			6	1						7	86%	6
fake (tease)	8					1				9	89%	8
simultaneous hit	8								1	9	89%	8
shoo						12			1	13	92%	12
hit ground	15	1								16	94%	15
dangle	17									17	100%	
duck lips open	11									11	100%	
food beg (oral)								5		5	100%	
play face	5									5	100%	
object on head	6									6	100%	
raise arm	9									9	100%	
raspberry face	5									5	100%	
roll on back	6									6	100%	
tandem walk				7						7	100%	
wave	17									17	100%	
somersault	4									4	100%	
swing	4									4	100%	
arms up	4									4	100%	

When the most common goal of a gesture represented at least 70% of the use of that gesture and was used at least three times more frequently than the next most common goal, it was deemed to be the goal of that gesture and is represented in Table 7 by a shaded cell. Cells marked with asterisks represent the most common observable goal for gestures that did not meet the criteria for use in pursuit of a single goal. In the three columns at the far right, the total number of cases of each gesture type with observable goals is mentioned, along with the percentage of the time that each gesture elicited its most common goal, and the ratio of the number of times the gesture was used towards the most common goal to the next most common goal. The gestures are ranked in Table 7 according to how often they were used towards a single goal. Most of the gestures were used in conjunction with the goals "affiliate/play," "move back/leave," and "stop." Gestures were not evenly distributed across observable goals, indicating that the categories of observable goals I created were indeed different for the orangutans.

3.4.3 Reconstructing gesture boundaries

In order to determine whether the gestures I had originally defined by form and use were meaningful signals to the orangutans, I compared the observable goals of each gesture type for the set of "successful" gestures to see how often they were used towards a single observable goal. I then redefined the boundaries for some gestures so that they were split into more specific gestures each representing as few goals as possible while still remaining frequent enough to allow for further analysis. I determined that a gesture type had a particular observed goal if 1) it was used more than 70% of the time towards one presumed goal, and 2) it was used more than three times as often towards that goal as it was towards any other goal. If the primary observable goal for a gesture reached the 70% and three-times as frequent as the secondary goal criteria, then that gesture type was considered to have a meaning corresponding to that goal.

Out of the reduced dataset of 38 gestures with frequent observable goals (Table 7), 27 were used predictably towards a single goal, and 11 were used towards more than one goal. For these 11 gesture types used towards several different observable goals, I examined additional structural and social variables to determine whether each gesture could be divided into two or more gestures, which were used at least 70% of the time to reach a particular goal and more than three times as often towards that goal as towards a secondary goal. This attempt to reduce the number of meanings for each gesture was used to best fit the observed movements to the set of observable goals. In addition, this approach allowed me to determine which structural or social variables were most important in distinguishing one gesture from another. The variables I analysed in my attempt to find parsimony in gesture and meaning are presented in Table 8.

Table 8: ADDITIONAL VARIABLES ANALYSED AS POTENTIALLY DEFINING CHARACTERISTICS TO REDUCE AMBIGUITY IN GESTURE.

Type of variable	Variable	Definition
Structural Variables	Facial expression	Signaller's facial expression
	Final hold	Final position of gesture is held (either tactile or visual)
	Handshape	Shape of hand in manual gestures (e.g. fist, flat hand, relaxed hand)
	Limb used	Part of the body performing the gesture (hand, foot, right vs. left)
	Target location	The location to which the signaller directs the gesture (either tactile or visual). It may be on the signaller or an object
Social Variables	Age class of signaller	Age of signaller (infant, juvenile, subabult, adult)
	Observers present	Whether the non-recipient individuals are present during the gesture
	Recipient's behaviour before gesture	The activity of type of behaviour the recipient is engaged in before the signaller's first gesture
	Recipient's visual attention	Whether the recipient can potentially see the signaller
	Relationship to recipient	The signaller's social relationship to the recipient based on age class or dominance (e.g. mother to infant, dominant to subordinate, infant to juvenile)

3.4.4 Analysis of additional variables

The 11 gestures that were habitually used towards more than one goal included the five most frequently observed potential gestures (*pull, touch, grab, push, grasp*). The 11 gestures analysed were (in ascending order of percentage use towards one goal): *grasp, pull, turn head, grab, mouth, touch, brush, swat, pull away, reach, push*. Both structural and social variables (Table 8) were examined to determine whether any of these gestures could be further divided to better predict differences in use.

Structural variables

The structural variables coded were 1) facial expression, 2) final hold, 3) handshape, 4) limb used, and 5) target location. Table 9 shows whether any of these variables were able to divide the original gestures into two or more parts, thereby creating any more specific gestures which met the criteria for use towards a single goal. Any proposed gesture based on commonalities along a structural variable had to have been observed at least 4 times towards one particular goal and had to meet the 70% prediction success and three times secondary goal criteria.

Table 9: SUCCESS OF ADDITIONAL STRUCTURAL VARIABLES IN REDUCING AMBIGUITY IN GESTURE MEANING.

The number of observed tokens of each gesture is presented in parentheses beside the gesture name (356 tokens were included in the table). The number of tokens that display a specific value of a variable is presented in parentheses, and the percentage of that number that have the same observable goal is given afterwards.

Gesture	***Facial Expression***	***Hold***	***Handshape***	***Limb***	***Target***	**Possible gestures?**
Grasp (43 / 35%)	—	—	—	—	food/ object (5) 100%	-***Grasp food/object*** -*Grasp*
Pull (78 / 37%)	—	—	—	—	head (15) 73% food/ object (18) 72%	-***Pull head*** -***Pull food/object*** -*Pull*
Turn head (13 / 38%)	—	—	—	—	—	-*Turn head*
Grab (48 / 46%)	—	—	—	both hands (6) 83%	food/ object (5) 80%	-***Grab (both hands)*** -***Grab food/object*** -*Grab*
Mouth (8 / 50%)	—	—	—	—	—	-*Mouth*
Touch (53 / 53%)	—	—	finger (5) 100%	—	—	-***Touch (finger)*** -*Touch*
Brush (12 / 58%)	—	—	—	—	—	-*Brush*
Swat (22 / 59%)	—	—	—	—	face/head (10) 80%	-***Swat face/head*** -*Swat*
Pull away (10 / 60%)	—	—	—	—	hand/foot (6) 83%	-***Pull away appendage*** -*Pull away*
Reach (25 / 64%)	—	hold position (13) 85%	—	—	—	-***Reach hold*** -*Reach*
Push (44 / 68%)	—	—	cupped or gripped (15) 87%	—	head region/ back (26) 81%	-***Push (grip)*** -***Push head region/back*** -*Push*

Some gestures can be categorised successfully according to more than one variable, so the numbers in the variable columns in Table 9 sometimes summed to more than the observed number of gestural tokens. The final column, “split?,” indicates what new gestures (if any) could be formed from each original gesture. The new gestures are proposed according to the variables that can explain the greatest amount of variance. The gestures written in bold in the "split?" column are those that were used predominately to achieve a single goal. When new gestures were proposed from several non-mutually exclusive variables, the new gestures that appear in the "split?" column are ranked according to their relative success in accounting for variance in observable goal. This ordering was to ensure that variables that often co-varied were not being used to artificially inflate the number of proposed new gestures. Though the numbers in each variable column represent the proportion of the total number of original gestures that fell into that category, the new gestures were formed each according to a single variable and once assigned to a new gesture, were removed from the original gesture dataset. Therefore, all possible new gestures are prioritized in terms of the variables and are non-overlapping.

Social variables

The social variables coded were 1) signaller's age class, 2) observers present, 3) recipient's behaviour before start of gestural interaction, 4) recipient's visual attention, and 5) signaller's relationship to recipient. Table 3 shows whether any of these variables were able to define a subset of examples of each gesture so as to accurately predict observable goal. As for the structural variables, any subdivided group defined by a common value of one of the social variables had to have been observed at least 4 times towards one particular goal.

Table 10: SUCCESS OF ADDITIONAL SOCIAL VARIABLES IN REDUCING AMBIGUITY IN GESTURE MEANING.

Gesture	**Signaller's age class**	**Observers present**	**Recipient's behaviour before gesture**	**Recipient's visual attention**	**Signaller's relationship to recipient**	**Possible splits?**
Grasp (43 / 35%)	—	—	—	—	—	*-Grasp*
Pull (78 / 37%)	—	—	foraging (9) 100% feeding (7) 100%	—	—	***-Pull (forager)*** ***-Pull (feeder)*** *-Pull*
Turn head (13 / 38%)	—	—	—	—	—	*-Turn head*
Grab (48 / 46%)	—	—	rest-watching (7) 71%	—	infant to immature (8) 100%	***-Grab (infant to immature)*** ***-Grab (rester)*** *-Grab*
Mouth (8 / 50%)	—	—	—	—	immature to adult (4) 100%	***-Mouth (immature to adult)*** *-Mouth*
Touch (53 / 53%)	—	—	locomoting (6) 83%	—	infant to non-mother (8) 88%	***-Touch (infant to non-mother)*** ***-Touch (locomotor)*** *-Touch*
Brush (12 / 58%)	—	—	attempting an interaction (7) 100%	looking (7) 86%	down hierarchy (age) (10) 70%	***-Brush (recipient attempting interaction)*** *-Brush*
Swat (22 / 59%)	immature (14) 86%	—	rest-watching (10) 90%	—	immature to adult (10) 90%	***-Swat (immature to adult)*** *-Swat*
Pull away (10 / 60%)	—	—	—	—	—	*-Pull away*
Reach (25 / 64%)	adult (11) 72%	—	rest-watching/ vigilance (10) 90%	—	—	***-Reach (watcher)*** *-Reach*
Push (44 / 68%)	adult (21) 90%	—	rest-watching/ vigilance (9) 78%	looking (30) 70%	down hierarchy (age/rank) (24) 92%	***-Push (down hierarchy)*** *-Push*

3.4.5 Index of effectiveness

Both structural and social variables were effective in predicting specific observable goals in gestures. The only variables that were never good predictors of observable goals were facial expression and the presence of observers. These variables appear to have no effect in determining meaning of gestures or were not used frequently enough to be detected as effective variables.

In order to determine which variable values were the best predictors of meaning for each gesture and across all gestures, I calculated an index of effectiveness for each variable. This index was calculated for each variable value that was found to be useful in predicting observable goal, and took into account the proportion of gestures of that type that fit the description as well as the variable's rate of success in predicting a single meaning. The goal of the index was to help determine which variable value was the best predictor of successful outcome (observable goal) for each gesture. This index of effectiveness was calculated by multiplying *Frequency of goal within variable* by *Prediction rate* for each variable value that successfully predicted a specific goal for each gesture. The index was calculated according to the following formula:

$$\mathcal{E}\text{index}_{[\text{ValueA,GoalJ,GestureX}]} = \textit{Frequency}_{[\text{GoalJ,VariableA}]} * \textit{Prediction rate}_{[\text{A}]}$$

More specifically:

$$\mathcal{E}\text{index}_{[\text{ValueA,GoalJ,GestureX}]} = \frac{\mathcal{F}_{[\text{A,J}]}}{\text{J}} * \frac{\mathcal{F}_{[\text{A,J}]}}{\mathcal{F}_{[\text{A}]}}$$

Where:

$\mathcal{E}\text{index}_{[\text{ValueA,GoalJ,GestureX}]}$ = Effectiveness of ValueA in predicting GoalJ for GestureX

J = number of tokens of Gesture X that have the observable Goal J

$\mathcal{F}_{[\text{A}]}$ = Number of tokens of GestureX that exhibit ValueA (regardless of goal)

$\mathcal{F}_{[\text{A,J}]}$ = Number of tokens of Gesture X exhibiting ValueA that have Goal J

To illustrate how this index would function, consider the following fictional example. Suppose we have a gesture "*shake hands*," which we have observed a total of 100 times. Out of those 100 examples, 50 were directed towards the goal "affiliate." (This percentage (50%) does not meet our threshold of 70%, so we do not consider it to have a consistent meaning based on our initial definition of *shake hands*.) We then examine other structural and social variables and discover that the structural variable "handshape," when it had the value "loose grip," was observed 60 times, and predicted the observable goal "affiliate" 45 out of 60 times (75% of the time). We also discover that when handshape had the value "tight grip," it predicted the goal "affiliation" 100% of the time, but was only observed 5 times. Finally, we found that the social variable "signaller's relationship to recipient," when it had the value "subordinate to dominant," was observed 40 times, and predicted the goal "affiliate" in 35 cases (88% of the time). So, we have one variable value (handshape:"loose grip") with 45 observations and a 75% prediction rate, one variable value (handshape:"tight grip") with 5 observations and a 100% success rate, and one variable value (social relationship:"subordinate to dominant") with 40 observations and an 88% prediction rate.

Shake hands	Handshape	Handshape	Relationship
100 total	"loose grip"	"tight grip"	"sub to dom"
(50 = affiliate)	60 total (45=aff.)	5 total (5=aff.)	40 total (35=aff.)

To determine which variable value is a better predictor of affiliate for the gesture *shake hands*, we calculate the index of effectiveness for each:

$$\mathcal{E}\text{index}_{(\text{loose grip,affiliate,shake hands})} = (45/50) * (45/60) = .9 * .75 = \mathbf{.68}$$

$$\mathcal{E}\text{index}_{(\text{tight grip, affiliate, shake hands})} = (5/50) * (5/5) = .1 * 1 = \mathbf{.1}$$

$$\mathcal{E}\text{index}_{(\text{sub to dom, affiliate, shake hands})} = (40/50) * (35/40) = .8 * .88 = \mathbf{.7}$$

Comparison of the three indices leads one to conclude that the relationship “subordinate to dominant” is the best predictor of the goal “affiliate” for the gesture “shake hands.” The handshape “tight grip” had the highest prediction rate, but explained very little of the variance in meaning for the total set of *shake hands*. The handshape “loose grip” was the most frequent value, but it was not as successful at predicting the goal “affiliate” as either of the two other variables. The comparison of indices prioritises the relationship “subordinate to dominant,” indicating that this variable value is the best predictor of meaning for the gesture *shake hands*.

This index provides a way to compare the effectiveness of different variables in predicting the observable goal (or “meaning”) of a gesture. If the values of the predictive variables are not mutually exclusive (as two handshapes would be), the variables may overlap and may covary a significant portion of the time (as handshape and relationship did in our example). However, if they do not covary 100% of the time, it is necessary to determine which variable is the best predictor of the variance of observable goal within all examples of a gesture type. Moreover, by determining which variables are the best predictors of meaning, it is possible to draw inferences about whether the meaning of gestures is determined more by structural or social variables. To put it another way, it allows us insight into whether the *gesture* or the *interaction* is more important in determining meaning. It may be the case that the structure of the gestures themselves carries meaning that others can use to determine the signallers goals. However, it may also be the case that the observers use information from the context of the gestures (e.g. the identity of the signaller, the observer’s own actions prior to the gesture, or their relationship to the signaller) in order to predict the signaller’s goal. Comparing the effectiveness of structural and social variables as predictors of the signaller’s goal brings us one step closer to understanding how orangutan gestural communication functions, whether it guides or reacts to social interaction.

For the 11 gestures that were used frequently to more than one goal and were subsequently analysed in greater detail, indices of effectiveness ($\mathcal{E}$index) were calculated when more than one variable value reliably predicted an observable goal. The indices were calculated for 9 of the gestures. The gesture *turn head* had no additional predictor of meaning, and the gesture *mouth* had only one, so these gestures were not given indices. The indices for the remaining 9 gestures are presented in Table 5. Each cell in the table presents a variable value, the goal it predicts, the number of gestures that fit the description, the number of those gestures that are used towards the goal, the total number of gestures that are used towards that goal, and the effectiveness index. The values are organised in each cell in the following manner:

Gesture X	variable:VALUEA ($\mathcal{F}_{[A]}$) *GOAL* [J] ($\mathcal{F}_{[A,J]}$)	$\mathcal{E}$index
Gesture X	variable:VALUEA (# gesture X observed with valueA) *GOAL* [# gesture X to *goal*] (# gesture X of valueA to *goal*)	Effectiveness index

Table 11: EFFECTIVENESS INDICES FOR ALL STRUCTURAL AND SOCIAL VARIABLES THAT PREDICTED AN OBSERVABLE GOAL FOR ONE OF THE 11 AMBIGUOUS GESTURES.

Total numbers of each gesture are given in parentheses in the left column (356 uses of gesture were analysed in all). Structural and social variables are grouped in columns and all significant variable values for a gesture are grouped in rows by gesture. Values used to calculate effectiveness indices for each gesture are presented in the following manner variable:VALUEA ($\mathcal{F}_{[A]}$) *GOAL* [J] ($\mathcal{F}_{[A,J]}$). The highest effectiveness index for each gesture is shaded in grey. When two indices are equal or when the variable values predict different goals for a gesture, both indices are shaded.

Gesture	Structural		Social	
Grasp (43)	target:FOOD/OBJECT (5) *SHARE* [5] (5)	1		
Pull (78)	target:HEAD (15) *FOLLOW* [29] 11	.3	recipient behav:FORAGE (9) *FOLLOW* [29] (9)	.3
	target:FOOD/OBJECT (18) *SHARE* [16] (13)	.6	recipient behav:FEED (7) *SHARE* [16] (7)	.4
Grab (48)	limb:BOTH HANDS (6) *AFFILIATE/PLAY* [22] (5)	.2	recipient behav:REST-WATCH (7) *AFFILIATE/PLAY* [22] (5)	.2
	target:FOOD/OBJECT (5) *SHARE* [7] (4)	.5	relationship:INFANT TO IMMATURE (8) *AFFILIATE/PLAY* [22] (8)	.4
Touch (53)	handshape:FINGER (5) *AFFILIATE/PLAY* [28] (5)	.2	recipient behav:LOCOMOTION (6) AFFILIATE/PLAY [28] (5)	.1
			relationship:INFANT TO NON-MOTHER (8) AFFILIATE/PLAY [28] (7)	.2
Brush (12)			recipient behav:ATTEMPT INTERACTION (7) STOP [7] (7)	1
			recip visual attention:LOOKING (7) *STOP* [7] (6)	.7
			relationship:DOWN HIERARCHY (10) *STOP* [7] (7)	.7
Swat (22)	target:FACE/HEAD (10) *AFFILIATE/PLAY* [13] (8)	.5	age class:IMMATURE (14) *AFFILIATE/PLAY* [13] (12)	.8
			recip visual attention:LOOKING (10) AFFILIATE/PLAY [13] (9)	.6
			relationship:IMMATURE TO ADULT (10) *AFFILIATE/PLAY* [13] (9)	.6
Pull away (10)	target:HAND/FOOT (6) *STOP* [6] (5)	.7		
Reach (25)	hold:HELD (13) *AFFILIATE/PLAY* [16] (11)	.6	age class:ADULT (11) *AFFILIATE/PLAY* [16] (8)	.4
			recip behaviour:REST OR STAND & WATCH (10) *AFFILIATE/PLAY* [16] (9)	.5
Push (44)	handshape:CUPPED/GRIPPED (15) *MOVE BACK/LEAVE* [30] (13)	.4	age class:ADULT (21) *MOVE BACK/LEAVE* [30] (19)	.6
	target:HEAD REGION (26) *MOVE BACK/LEAVE* [30] (21)	.6	recip behaviour:REST OR STAND & WATCH (9) *MOVE BACK/LEAVE* [30] (7)	.2
			recip visual attention:LOOKING (30) *MOVE BACK/LEAVE* [30] (21)	.5
			relationship:DOWN HIERARCHY (24) *MOVE BACK/LEAVE* [30] (22)	.7

3.4.6 Meanings

After comparing the effectiveness indices of the structural and social variables for the 9 gestures, I identified the variable that best predicted meaning in each gesture. Structural variables most effectively predicted observable goal in the gestures *Grasp*, *Pull*, *Pull away*, and *Reach*. Social variables most effectively predicted observable goal for the gestures *Touch*, *Brush*, *Swat*, and *Push*. The observable goal of the gesture *Grab* was best predicted by a structural variable when the goal was "share," and by a social variable when the goal was "affiliate/play." The gesture *Push* was also predicted to have two different goals once the additional variables were examined, but both goals were predicted by structural variables. Overall, structural variables were better predictors of observable goal than social variables were.

Of the 6 cases where structural variables best predicted the observable goal of a gesture, 4 were predicted by "target location," 1 by "handshape," and 1 by "hold." Of the 5 cases of social variables most strongly predicting an observable goal, "relationship between signaller and recipient" was the most effective variable 3 times, "recipient's behaviour" and "signaller's age class" was each the most effective variable once. It seems clear that the target location of the gesture and the relationship between signaller and recipient are important variables in predicting the observable goal of gestures. This comparison is muddied slightly by the fact that the target location "food/object" accounted for 3 of the 4 cases where target was the best predictor of meaning. In all of these 3 cases, the observable goal was "share." It seems reasonable that the target location would be more important than other gestural variables in predicting that specific goal. Perhaps all gestures that touch an object held by another should be defined as a single gesture by their target location rather than grouped with other gestures through similarity of movement.

The comparison of the effectiveness index illustrates how both structural and social variables combine to create meaning and predict outcome. The structural variables "target" and "handshape" were both frequently able to predict the observable goals of gestures. The social variables “age of signaller” and “signaller’s relationship to recipient” were frequent predictors of observable goals as well. The recipient’s behaviour before the interaction was also able to control for some of the variance in meaning. Though it was only the most effective variable in predicting meaning once, the recipient’s behaviour before the gesture could be used to accurately predict an observable goal in 8 out of the 11 gestures analysed. This made it the most common, though not the most effective, predictor of meaning.

To most accurately interpret another’s meaning or to choose how to communicate most effectively, an orangutan must take into account both the form of the gesture and the nature of the preceding social interaction. I suspect that orangutans are monitoring both structural and social elements and are fairly sophisticated in their communicative strategies. Many of their gestures accurately predict one outcome when they are successful (i.e. when the presumed goal and outcome match).

In all, 6 different meanings were reliably predicted by gestures, and an additional meaning was predicted once further variables had been considered. Table 12 gives the gesture types associated with each meaning. The goals “pay attention to” and “pick up” were not accurately predicted by any gesture. It is likely that either there is no discernable difference between “pay attention to” and “affiliate/play,” or that the observers were not able to distinguish between the two. It seems reasonable that “pay attention to” should be added to “affiliate/play” rather than being counted as a separate goal.

Table 12: OBSERVABLE GOALS ACCURATELY PREDICTED BY SPECIFIC GESTURES.

The gestures without shading were those that met the criteria for 70% use towards a single goal that was 3 times more common than the secondary goal. They are listed in order of total number of observations toward that goal. Gestures in shaded cells were ambiguous gestures for which a subset could be shown to accurately predict a specific goal once further variables had been considered. The “numbers accurately predicted” for the shaded gestures were those that increased the accuracy rate to more than 70% when one of the additional variables was taken into account. The column on the far right lists the additional variables that accurately predict specific goals for the ambiguous gestures. The table represents analysis of 639 gestural tokens.

Meaning (observable goal)	Gesture	Number accurately predicted	Additional variables for prediction?
Affiliate/Play	Hit	28	
	Dangle	17	
	Wave	17	
	Hit ground	15	
	Bite	13	
	Pull hair	11	
	Duck lips open	11	
	Raise arm	9	
	Embrace	8	
	Fake	8	
	Simultaneous hit	8	
	Put object on head	6	
	Roll of back	6	
	Air bite	5	
	Play face	5	
	Raspberry face	5	
	Back roll	4	
	Shake	4	
	Arms up	4	
	Touch	7, 5, 5	Relationship, Handshape, Recipient behaviour
	Grab	8, 5, 5	Relationship, Limb, Recipient behaviour
	Reach	11, 9, 8	Hold, Recipient behaviour, Age class
	Swat	12, 9, 9, 8	Age class, Recipient visual attention, Relationship, Target
Move back/Leave	Nudge	19	
	Shoo	12	
	Push	22, 21, 19, 21, 13, 7	Relationship, Target, Age class, Recipient visual attention, Handshape, Recipient behaviour
Share food/object	Food beg orally	5	
	Pull	13, 7	Target, Recipient behaviour
	Grab	4	Target
	Grasp	5	Target
Follow	Tandem walk	7	
	Pull	11, 9	Target, Recipient behaviour
Climb on	Embrace pull	6	
Look at object	Offer	5	
Stop	Brush	7, 7, 6	Recipient behaviour, Relationship, Recipient visual attention
	Pull away	5	Target

Affiliation/play was far and away the most frequent goal and was associated with the greatest number of gesture types. Some of the gestures aimed at initiating affiliation or play are idiosyncratic (see 3.6.1) or used by only a few individuals and so the set of gestures used by most orangutans to initiate affiliation or play is only a subset of those listed in Table 12. Also, the goals “affiliation” and “play” were combined into a single goal because it was difficult to distinguish low-level play (i.e. not wrestling) from affiliation, and gestures associated frequently with one goal were also associated frequently with the other. The category represents a wide range of possible outcomes: from highly active play (wrestling), to sitting in contact, to manipulating objects together.

3.4.7 Evaluation of gesture definitions

Although 10 goal-specific gestures could be created as subsets of 9 of the 11 ambiguous gestures once additional structural and social variables were taken into account, they should not all be classified as new gestures. If, by considering an additional variable, an ambiguous gesture could be split into two single-goal gestures, it is reasonable that they should be considered separate, regardless of whether there are only a few examples. However, most of the ambiguous gestures I examined were not divided into two single-goal gestures, but rather managed to have only a subset of the gesture meet the threshold for single-goal use by defining the gesture more narrowly. This narrow definition drastically reduced the number of examples of each gesture.

When there is a clear case that additional variables can account for the ambiguity of a gesture by cleanly dividing it into two or more narrowly-defined gestures, it becomes clear that what was initially considered to be a single gesture was actually several gestures. However, most of the variables that I identified as playing a role in predicting meaning in ambiguous gestures were, in fact defining only a small subset of the total examples of the gesture, and only rarely predicted a single outcome with high accuracy. This selective whittling down of gestures would result in an over-specificity of the new gestures that would render many further analyses impossible by increasing the number of rare gestures to the point where comparison across individuals or contexts became insignificant.

Social variables such as relationship between signaller and recipient age clearly play a role in shaping the meaning of the gesture, but it is possible that these social factors do not influence the gesture as much as the goal of the signaller. The signaller's goals may change with age or with their relationship to other individuals. It is likely that the gestures reflect different goals not because the gestures themselves change, but because they are performed by a different type of individual. Any discussion of meaning of communicative signals assumes that the signal itself carries meaning. However, it is clear that contextual variables such as age and the recipient's actions heavily influence the success and outcome of a particular gesture. It may well be the case that to properly interpret the goal of the signaller or to predict how the recipient will respond, one needs to understand the context in which the gesture occurs. The question this raises is: does the context affect the gesture itself?

In English, words or gestures sometimes have different meanings depending on context. This might be due to metaphorical use or common properties in both referents (e.g. "flip" as an action versus a hairstyle). However, in these cases, the meaning is determined by a comparison of the word and context, not because the word itself has changed. I would also argue that a gesture should not be defined as a new type when it is used within a different social situation. It seems to me that differences in meaning that spring from social variables should be seen as part of the flexibility of the gesture rather than as a different gesture altogether. If gestures are defined by social relationships, then not all individuals have the opportunity to learn, develop, or use all gestures. It is counterintuitive to define a set of intentional, communicative signals that cannot be performed by certain individuals. The goal of developing a species repertoire of gestures is to compile a set of all possible gestures for the species. If social variables are used to define gestures, individuals of certain ages or ranks will need to have different potential repertoire sizes by definition.

The fact that the measured variable used to determine meaning was "observed goal," and restricted the dataset to tokens in which the outcome met the presumed goal of the signaller, had an effect on the types of gestures that were included in the analysis. The analysed set of gestures represented only "effective" or "successful" gestures (those that immediately or eventually received the desired response). It is likely that social variables have an influence over which gestures are successful as older individuals may be more likely to ignore gestures of infants or juveniles or dominant individuals may be less likely to response to gestures by subordinates. These effects would have a great impact on the number of gestures with observable goals. Immatures (infants and juveniles) gesturing to adults receive no response 28% of the time. Conversely, adults gesturing to immatures receive no response only 13% of the time. Infants therefore "fail" in their gestures more than twice as often as adults. The clear effect of relationship between signaller and recipient on the success of gestures to receive a response suggests that the set of gestures with "observable goals" may be skewed towards gestures used in certain social situations. Most of the gestures that did not meet the criteria for an observable goal had presumed goals, but the reactions of the recipients failed to meet them. Instances where social variables such as the relationship between signaller and recipient can predict observable goals in a subset of a gesture, may, in fact, be more a reflection of the relative success in achieving a desired response than a real change in meaning. For this reason, no gestures were redefined using social variables. The analysis of social variables was used to inform our understanding of the role contextual use can play in shaping meaning, but was not used to create new gestures.

The strongest case for the creation of a new gesture is the convergence of meaning of *grasp*, *grab*, and *pull* when directed towards an object held by the recipient. I believe that this is an ideal case where comparison of the contextual use and recipient's response indicates that these gestures are used interchangeably under the same circumstances. The three gestures should be combined into one that prioritises the target location as a defining variable over the variables of movement, speed, and force that were initially used to divide gestures. The target location "food/object" seems to only be effective in predicting the goal "share" in the three gestures *grasp*, *pull*, and *touch*. As pulling, grasping, and grabbing an object all seem to have the same observable goal, but are distinct from other gestures directed towards objects, I combined the three gestures (when object-directed) into the single gesture "*seize object*." All tokens of *pull, grab,* and *grasp* that were aimed at objects were therefore renamed, regardless of whether they had an observable goal.

Structural variables were only used to define new gestures if they received an index of effectiveness of .7 or higher. This threshold reduced the number of new gestures formed and ensured that each new gesture is not only internally consistent (i.e. directed towards one goal), but also accounts for a significant amount of the variance of meaning within the original gesture. The only gesture to meet this criterion for subdivision into a new gesture is *pull away*. When *pull away* is directed at the recipient's hand or foot, it predicts the observable goal "stop" with an effectiveness index of .7. Thus, I divided *pull away* into two gestures: *pull away* and *pull away (appendage)*.

Low-fidelity gestures

The ambiguous gestures that successfully predicted single goals with the addition of structural variables but received effectiveness indices of less that 0.7 were not subdivided into new gestures as the additional variables did not predict enough of the variance of meaning. However, rather than assume that all ambiguous gestures did not have goals, I defined a new category of "low-fidelity" gestures. This category included all gestures that were used towards one goal 50%-70% of the time at a rate at least twice that of the secondary goal. This category includes the gestures *grab, push, reach, swat, touch, push, mouth,* and *brush* (Table 13).

Table 13: HIGH AND LOW FIDELITY GESTURES FOR EACH OBSERVABLE GOAL.

The high-fidelity gesture types listed accounted for 320 tokens of gesture, while the low-fidelity gesture types accounted for 204 of the total set of 698 gestural tokens with observable goals.

Observable goal	**High Fidelity Gestures (70% and 3 times rate of secondary goal)**	**Low Fidelity Gestures (50%-69% and 2 times rate of secondary goal)**
Affiliate/Play	*-Air bite* *-Arms up* *-Back roll* *-Bite* *-Dangle* *-Duck lips open* *-Embrace* *-Fake* *-Hit* *-Hit ground/object* *-Play face* *-Pull hair* *-Put object on head* *-Raise arm* *-Raspberry face* *-Roll on back* *-Shake object* *-Simultaneous hit* *-Somersault* *-Swing* *-Wave*	*-Grab* *-Reach* *-Swat* *-Touch*
Move back/Leave	*-Nudge* *-Shoo*	*-Push*
Share	*-Food beg orally* *-Seize*	*-Mouth*
Stop	*-Pull away appendage*	*-Brush*
Follow	*-Tandem walk*	
Look at object	*-Offer*	
Climb on	*-Embrace pull*	

3.5 Do some gestures have different meanings for different individuals?

Following the initial analysis of meaning based on how often gestures correlated with observable goals, I investigated the possibility that the observed variance in meaning might be caused by individuals who used the gestures solely towards a different goal. In this scenario, some orangutans might come to associate gesture *x* with meaning x, whereas others might associate it with meaning y. This proposed variability of meaning would indicate that, while the forms of most gestures might be species-typical and likely heritable, the meanings of the movements are not fixed but must be learned through association with response from different recipients (as development by ontogenetic ritualisation would predict, see Call and Tomasello 2007). If, on the other hand, variability in meaning is found to be the result of occasional variation in use by all individuals and no individual uses only the secondary meaning, then there is no support for the theory that meaning is individually learned during ontogeny.

The secondary goals of the 29 high-fidelity gestures and the 7 low-fidelity gestures were examined if at least 2 examples of a gesture were observed towards the secondary goal (Table 14). The aim was to determine whether the instances of gestures performed towards secondary goals were the product of occasional variation in use or because some individuals used only the second meaning of the gesture. Twenty-five gestures (*air bite, arms up, back roll, dangle, duck lips open, embrace, fake, hit ground/object, play face, pull hair, put object on head, raise arm, raspberry face, roll on back, shake object, simultaneous hit, somersault, swing, wave, shoo, food beg orally, pull away appendage, tandem walk, offer,* and *embrace pull)* were observed fewer than 2 times in conjunction with any secondary goal and so were not used in analysis of secondary meaning.

Table 14: PRIMARY AND SECONDARY MEANINGS OF GESTURES.
The table presents meaningful gestures (as defined in 3.4) along with their primary and secondary meanings (observable goals). Only gestures that were observed at least twice in conjunction with a secondary meaning were included. The numbers of individuals that exhibited the primary and secondary meanings of the gestures are listed along with the ratio of observations of the 1st and 2nd meanings for each individual that was observed using the 2nd meaning. The final column lists any commonalities amongst individuals that used the gesture more often in conjunction with the 2nd goal.

Primary meaning	***Gesture***	**Secondary meaning**	**# of indv. using 1st meaning**	**# of indv. using 2nd meaning**	**Observations of 1st:2nd meanings for those indv.**	**Explanation for indvs. with 2nd meaning**
Affiliate /Play	*Bite*	Stop	7	3	0:1, 0:1, 0:1	Adults
	Hit	Stop	11	3	4:1, 2:4, 0:1	Adults
	Grab	Stop	13	5	6:1, 1:3, 0:1, 0:2, 0:3	Adults
	Reach	Pick up	10	1	1:3	Infant
	Swat	Move back/ Leave	5	4	3:1, 0:1, 0:2, 0:2	Adults
	Touch	Move back/ Leave	13	6	0:1, 0:1, 0:1, 0:1, 0:1, 0:2	Adults
Move back /Leave	*Nudge*	Stop	9	4	3:1, 1:1, 1:2, 0:1	"Move back" and "stop" easily confused by observer?
	Push	Stop	12	5	1:1, 1:2, 0:1, 0:1, 0:2	"Move back" and "stop" easily confused by observer?
Share	*Seize*	Affiliate/ Play	6	2	8:2, 0:2	Both tokens from same interaction
	Mouth	Affiliate/ Play	2	2	3:2, 0:1	
Stop	*Brush*	Move back/ Leave	6	2	1:1, 1:1	

Table 14 displays the frequency of the secondary meanings for 11 gestures that I determined accurately predicted a single goal the majority of the time. The gestures in shaded cells are low-fidelity gestures that were used between 50% and 69% of the time towards a single goal. The unshaded gestures were used at least 70% of the times towards the same goal. The individuals using the first meaning and the number using the second meaning usually overlap as some individuals were observed using both meanings. Some gestures had tertiary meanings as well, but only primary and secondary meanings were analysed as the number of examples for tertiary meanings (and beyond) dwindled quickly. The ratio of examples of the primary and secondary meanings was listed for each individual who used the secondary meaning of a gesture.

Five gestures (*bite, hit, grab, swat, touch*) used primarily towards the goal "affiliate/play" had secondary meanings of either "move back/leave" or "stop." For these 5 gestures, the secondary meaning was used primarily by adults, while the primary meaning was used most frequently by immatures. Four of these gestures (*bite, hit, grab,* and *swat*), are energetic, aggressive movements, which are used frequently by immatures during play, but which do not feature as prominently in the affiliative interactions of older individuals. Instead, adults use these behaviours as serious threats or reprimands. It is possible that the difference in meaning associated with these gestures is an effect of infants using agonistic adult behaviours during play. Infant and juvenile orangutans often engage in rowdy wrestling play accompanied by biting and hitting. Juveniles of many species use adult behaviour during play (see Bekoff and Byers 1998). Perhaps immature orangutans have not yet come to use the gestures with agonistic goals, as others do not yet respond with subordinate actions. It is important to remember that the examples of gesture used to generate the meanings were all "successful" gestures where the outcome matched the presumed goal. As this definition relies on a gesture's being effective in order to be included in the analysis, it is possible that immatures do use these gestures to attempt to repulse other individuals but are unable to achieve their goals because of their lack of dominance.

The fifth “affiliate/play” gesture (*touch*) to have an agonistic secondary meaning was also used primarily by adults when employed towards its secondary goal. However, *touch* was the most ambiguous of the five as it was used frequently in conjunction with 6 of the 9 possible observable goals (Table 7). Two of the individuals who used *touch* with its secondary goal “move back/leave” used the gesture towards another non-primary goal as well, and none of the others were observed using *touch* more than once (with an observable goal). Because of this extreme ambiguity, this gesture is less likely to have an adult meaning and a juvenile one, and might instead be considered to be a truly ambiguous “catch-all” gesture. If this is the case, then the observed meaning “affiliate/play” is likely a product of the high frequency of the goal rather than a true meaning of the gesture. It seems that *touch* is more ambiguous than most gestures and most likely derives its meaning (when one is clear) from its context rather than its form or the status of its signaller. This extreme flexibility and reliance of contextual cues for meaning is very similar to deictic gesture (such as pointing) in humans (McNeill 1992). It is possible that touching in orangutans is similar to pointing in humans as it draws attention to the signaller’s intent to communicate something in a particular context rather than transmitting semantic information removed from contextual clues. If this is the case, then touching may be related to the “contact gestures” produced by gorillas (Gómez et al. 1993), but the use of touching by orangutans seems to be much less directed and specific. Is it is unlikely that touching is used to direct the attention of others to external entities, as pointing is in human infants (Liszkowski et al. 2004).

One infant used *reach* towards his mother when he wanted to be carried. He was also observed using *reach* to mean “affiliate/play,” but used the gesture most commonly used prior to being picked up. It is possible that this infant learned to associate the gesture with both meanings, or that his mother responded most frequently to his affiliative gesture by picking him up. Since he did not use the gesture exclusively with the observable goal “pick up,” and there are not enough examples of the gesture to analyse further variables, there is not enough evidence to conclude that this infant learned a different meaning for a common gesture. The infant’s mother was the most “worried” of all mothers in the study—she kept her infant close and nervously slapped him as he was held against her chest when the adult male approached. It is likely that this mother reacted more immediately to requests for affiliation or carrying from her infant, possibly resulting in the infant having increased success of gestures with the goal “pick up.” Without further examples of this gesture from this infant, I am unable to draw any conclusions as to the ontogeny of meaning.

The two gestures that were used towards the primary goal of “move back/leave” (*nudge* and *push*) both had “stop” as their secondary goal. They displayed similar ratios of use between the primary and secondary meanings and had no apparent common traits of individuals who used the gestures more often towards the secondary meaning. I suspect that the ambiguity present in both of these gestures is a product of experimenter error rather than a change in meaning. The goals “move back/leave” and “stop” can be highly similar, particularly when a gesture is directed at an individual who is attempting to affiliate with the signaller. It is difficult to discern if an adult signaller wants the infant that is pulling on her arm to go away or merely to let go; and, indeed, the orangutan herself might not distinguish between the two goals in this case. However, “move back/leave” would necessarily be the goal during acts of displacement, while the goal “stop” seems more likely to occur in situations where one individual attempts to dissuade rather than displace another. I suspect that these goals are distinguishable in some cases and not distinguishable in others leading to the observed ambiguity in meaning.

The secondary meaning of *seize* was “affiliate/play” and, unsurprisingly, the observed use of the secondary meaning occurred during object play such as tug-of-war. Since one of the two individuals who used the secondary meaning used the primary meaning more often, and the 2 instances in which another individual used the secondary meaning came from the same play session, I do not think that the gesture has a different meaning for these 2 individuals.

The final two gestures, *mouth* and *brush*, were not used frequently enough towards the secondary goal to draw any conclusions. *Mouth* was used towards its secondary goal “affiliate/play” by two individuals, one of who used *mouth* more frequently towards its primary goal and the other of whom was only observed to use it once (towards the secondary goal). *Brush* was also used by two individuals in conjunction with its secondary meaning, “move back/leave.” Both individuals used the gesture with its primary and secondary meanings, so the presence of the secondary meaning could result from true ambiguity of the gesture or, as was proposed for *nudge* and *push*, the discrepancy could be a product of experimenter error in distinguishing the two goals.

From this comparison of primary and secondary meanings, there is no good evidence that meanings are culturally transmitted. It appears that when gestures are used in conjunction with a secondary meaning, it is not because some individuals have learned a different meaning for the gesture, but rather a result of true ambiguity of meaning, or a reflection of changing social motivation. Social variables play an essential role in determining the meaning of gestures. Section 3.6 in this charter explores the effect of the signallers’ age on the gestures they use. Variables such as the age of the signaller or the relationship between signaller and recipient can lead to differences in meaning for the same gesture. This change reflects both changing social goals and effectiveness at achieving different results. The meaning of some gestures does differ between individuals, but it appears to do so generally as orangutans age, rather than as a product of cultural context. Meaning, it seems, is as universal as form in orangutan gestures.

3.6 Gesture use

In order to evaluate the use of different gestures, I included all observed gestures in the analysis, not only those with observable goals. The dataset of gestures with observable goals is particularly useful in determining meaning of different gesture types, but it cuts the number of observed gestures in half, reducing much of the variability in the data. In evaluating gesture use, I analysed all 1334 gestural tokens included in the dataset of intentional gestures (see 3.2). This dataset excluded gestures for each individual that had not been observed to use them in a manner indicative of intentional usage. The goal, observable or otherwise, of the gesture was not taken into account. Through analysis of this intentional dataset, I aimed to determine which individuals were using each gesture and whether any of the gestures were idiosyncratic, dyadic, group-specific, or species-typical. I was also able to investigate gesture use by different age classes (see 3.7.2).

3.6.1 Idiosyncratic gestures

For five of the gestures, no more than one individual was observed to use them intentionally. I therefore deemed these five gestures to be idiosyncratic, as only single individuals used the actions as communicative signals though others may have used similar movements in non-communicative ways. Some of these individuals used their idiosyncratic gestures to only a single recipient, others to many individuals in their group. The effectiveness of the gestures (whether or not they elicited the "intended" response) was not taken into account in determining whether these idiosyncratic gestures should be included as gestures; if they were accompanied by measures of intentional usage, they were counted as gestures, regardless of whether they ever received a response from their recipients. The five idiosyncratic gestures are listed in Table 15 along with the age class and zoo of the individuals that used them.

Table 15: "IDIOSYNCRATIC" GESTURES: GESTURES USED BY ONLY ONE INDIVIDUAL IN AN INTENTIONAL MANNER.

Gesture	Count	Modality	Group	Age	Sex
Drag	4	Visual	Durrell	Infant	M
Duck-lips (open)	11	Visual	Durrell	Adult	M
Fake	9	Visual	Twycross	Adult	F
Restrain	3	Tactile	Apenheul	Adult	F
Shrug	7	Tactile	Twycross	Adult	F

As idiosyncratic gestures were observed in all three zoos, no one group seemed to have a greater propensity for developing idiosyncratic gestures.

The infant that exhibited the gesture *drag*, used it in dyadic contexts of attempted play or in social contexts when the attention of the others was directed elsewhere. Thus, I assume that the infant used the action as an invitation to play. The gesture was only observed a total of four times in this individual, and never achieved a response from either of the two recipients to which it was directed. It is possible that the infant was either testing out this action as a gesture for the first time or that the action had succeeded before as a gesture and the lack of response in the four observations was an abnormality. Our data do not favour one interpretation over the other. It would be necessary to see whether the infant exhibited the gesture later in life in order to determine whether it had been established as a gesture or was observed by chance being used as a gesture. Either way, the infant used *drag* only rarely and had many other play initiation gestures. It is likely that *drag* was an object-directed play behaviour that could be used as a gesture towards another individual when coupled with response waiting or other gestures.

Duck-lips-open was used by a single adult male towards one of his infant sons to initiate play. The male used the gesture 11 times, and only once did it fail to receive a positive response. I concluded that this gesture was a highly effective invitation to play when used by this male to this infant.

Fake was used by an adult female in teasing contexts with her juvenile daughter. The teasing sessions appeared to take the form of a game where the adult female would quickly strike out towards a limb of the juvenile who would quickly retract it. Then the juvenile would slowly extend the limb again and the “game” would be repeated several times. I observed the same pair engaged in this routine several times during the study. One more than one occasion, when the juvenile turned away, the mother would use a tactile gesture to regain her attention before attempting *fake* again.

Restrain was used by an adult female at Apenheul Primate Park to cause others to stop their actions. It was a potentially physically-effective action in that the female would grasp the hand of another and hold it immobile. This action did not, however, hinder the other movements of the animal and the female never exaggerated this gesture by seizing other parts of the recipient's body. It seems likely that this behaviour is therefore an intentional gesture more than an effective action itself as it did not fully restrict the movements of the recipient but often caused him to cease what he was doing.

Shrug was used reliably by one adult female to either of her two daughters when they were grooming or attempting to affiliate with her. I interpreted the goal of the action as a desire that the recipient to back away, as the female repeated the gesture or switched to a more specific directional gesture such as a push if the *shrug* received no response.

Each of the five idiosyncratic gestures was used reliably and intentionally by a single individual. The gestures also seemed to be used to achieve specific goals as each was employed in a particular context. These idiosyncratic gestures seemed to be specialized gestures rather than generalist gestures; used for communicating one type of desire and, in two cases, to a specific individual. They were likely the effect of one individual finding that a specific movement or action elicited a desired response in another and then proceeding to use the action (or a ritualized version of the action) as a signal to others. Four of the gestures (*duck-lips-open*, *fake*, *restrain*, and *shrug*) were very effective in eliciting the intended reaction from the recipient. *Drag* did not appear to reliably elicit a particular reaction and it is possible that the infant who used this gesture was either "testing" the movement as a gesture, or particularly enjoyed the behaviour and persisted in using it to try to communicate despite continued failure.

3.6.2 Species-typical gestures

Thirty-five gestures were found to be present in every zoo (Table 16). These gestures were deemed to be “species-typical” as they were observed in all 3 populations. Of these, 18 were visual gestures and 17 were tactile gestures. All species-typical gestures were manual, oral, or whole-body gestures. No facial expressions were used as intentional gestures in all zoos. Though all thirty-five gestures were observed at each zoo, they differed in how common they were within each population. Table 16 shows the number of individuals at each zoo that exhibited intentional usage of the gestures.

Table 16: GESTURES PRESENT AT ALL SITES.

The number of individuals at each site that displayed each gesture is shown. The bracketed numbers beside the column heading for each zoo are the total number of animals at that site. The meanings were taken from the analysis in section 3.4.4.

Modality	Gesture	Apenheul (11)	Durrell (8)	Twycross (8)	Meaning
VISUAL	*Air bite*	4	1	3	Affiliate/Play
	Air grab	5	1	2	—
	Back Roll	1	1	1	Affiliate/Play
	Cover	1	1	1	—
	Dangle	4	5	1	Affiliate/Play
	Food beg orally	3	1	2	Share
	Hit ground/object	5	2	5	Affiliate/Play
	Offer	2	3	1	Look at object
	Peer	5	3	2	—
	Put object on head	2	1	1	Affiliate/Play
	Raise arm	5	2	3	Affiliate/Play
	Reach	5	6	6	Affiliate/Play (low fidelity)
	Roll on back	1	3	3	Affiliate/Play
	Shake object	3	1	2	Affiliate/Play
	Shoo	3	3	2	Move back/Leave
	Swat	5	2	5	Affiliate/Play (low fidelity)
	Swing	3	1	2	Affiliate/Play
	Wave	6	2	3	Affiliate/Play
TACTILE	*Bite*	6	3	4	Affiliate/Play
	Brush	6	2	5	Stop (low fidelity)
	Embrace	3	2	4	Affiliate/Play
	Embrace pull (guide)	1	1	1	Climb on
	Grab	7	5	6	Multi-goal
	Grasp	8	5	6	Multi-goal
	Hit	6	4	4	Affiliate/Play
	Nudge	5	4	4	Move back/Leave
	Poke	1	1	1	—
	Pull	8	5	5	Multi-goal
	Pull away appendage	4	2	2	Stop
	Pull hair	5	2	4	Affiliate/Play
	Push	7	5	7	Move back/Leave (low fidelity)
	Seize	3	4	4	Share
	Tap	3	4	4	—
	Touch	9	7	8	Affiliate/Play (low fidelity)
	Turn head	5	2	2	Multi-goal

Four of the gestures found in all three locations (*back roll*, *cover*, *embrace-pull* and *poke*) were only observed in a single individual in each zoo. However, these four gestures were limited to signallers within a specific age range (see section 3.7.2), and so the chance of observing them in multiple individuals at the same location is limited by the number of animals within the age range for potential signallers.

3.6.3 Group-specific gestures

Gestures that are found to be limited to a specific group can be evidence for local traditions within those groups, and for social learning of gesture. These types of gesture seem to be very rare in all ape species (see Call and Tomasello 2007), thus strengthening either the argument that gestures are ritualized by the same process regardless of which population an individual is in or that gestures are heritable movements shared by all members of a species that become used as gestures once recipients learn to predict the signaller's subsequent behaviour from their movements. Group-specific gestures can be an important tool in determining how movements become used as gestures as they shed light onto whether social learning can play a factor in shaping the development or expression of certain gestures.

In this study, four gestures were used intentionally by more than one individual in the same population, but not by any individuals in other zoos (Table 17). These gestures were deemed to be "group-specific." All four of these gestures are unusual, however, and each must be examined carefully in order to determine whether it is indeed a group specific gesture. All four gestures were each found in only two individuals in their group. This means that none of these "group-specific" gestures are common, even within a single population. Their rarity increases the likelihood that they exist in other populations but were never observed in an overtly intentional usage.

The likelihood of the four observed gestures' being truly limited to one group is complicated by the fact that three of the gestures are facial expressions. I considered facial expressions to be potential gestures if they were initiated and ended while the signaller was facing the recipient and made no other movements during the expression. They were then subjected to the same criteria of intentional use as all gestures made with other parts of the body. Three facial expressions that met the criteria for intentional gestures were found only in one zoo. Each of these expressions was observed in at least one other zoo as a facial expression accompanying a manual gesture but were never observed on their own. The expressions were only used as gestures unaccompanied by other movement in a single population each. It is possible that two individuals in a population found that a particular expression could communicate their desires on its own as well as when it accompanied a manual gesture.

Table 17: GROUP-SPECIFIC GESTURES.

The modalities of each gesture are given as well as the group in which they were observed. None of the gestures was observed frequently enough to be analysed for meaning.

Gesture	Count	Modality	Group	Meaning
Lip smile (tense)	9	Visual (facial expression)	Apenheul	—
Pout	5	Visual (facial expression)	Twycross	—
Tongue out	4	Visual (facial expression)	Twycross	—
Tandem walk	18	Tactile	Twycross	—

The manual gesture *tandem-walk* is an unusual gesture as it has been observed in all three populations but only seemed to be used as an intentional gesture in one of them. The action involves the signaller placing his hand on the shoulder or back of the recipient during side-by-side locomotion. The action often does not have an obvious goal, and instead may arise during an immature's transitional stage from being carried to walking independently. Alternatively, it could be used to guide the recipient during locomotion and to indicate a common destination. When *tandem-walk* met the criteria for intentionality, the action was accompanied by looking towards the recipient's face, and was alternated with pushing or pulling if the recipient did not move. The gesture was used intentionally and often by one juvenile (5.5-years-old) to guide his mother or position her between him and the adult male. One adult female in the same group used it in a similar way to guide her infant. Individuals (particularly immatures) at other zoos were observed placing a hand upon another's back during side-by-side walking, but the touch was not accompanied with any indication of intentionally communicative use. It may be the case that two animals at Twycross have begun to use a natural reaction to tandem locomotion as a signal; however, it is also possible that this "gesture" is a reactive and not communicative action and the examples of intentional use are unusual examples.

None of the gestures observed in only one zoo make a good case for local traditions or culturally-transmitted forms. The facial expressions are unique to one zoo only in being used without an accompanying gesture, and the low numbers of observations of the three expressions indicate that they are only rarely used on their own. The fact that each gesture was only observed in two individuals does not make a strong case for different cultural traditions at the different zoos. If the two signallers for each gesture used the gesture to communicate only with the other, one could argue that the actions had taken on meaning as gestures within certain dyads. However, none of the gestures were confined to a single dyad, though each were used by only two signallers. Most problematic is the fact that each of the actions was observed in other zoos, just not in a manner that fit my criteria for intentional usage. The use of the movements in other zoos, coupled with the infrequent use of three of the four gestures, leaves open the possibility that the actions were used as intentional gestures in the other zoos but were not observed frequently enough to be detected by the observer.

3.6.4 Gestures used in two of the three zoos

Twenty gestures were observed in 2 of the 3 zoos (Table 18). These gestures were mostly low frequency gestures; the mean number of tokens for the 20 gestures was 7.7 (± 8.9). Only three of these gestures, *arms up*, *mouth*, and *somersault*, were observed more than 10 times during the study.

Table 18: GESTURES OBSERVED IN 2 OF THE 3 ZOOS.

The gestures are presented with their modalities and meanings. Only 5 of the gestures were used consistently with one observable goal, and many were used too infrequently to be included in the analysis of meaning.

Gesture	Count	Modality	Group	Meaning
Arms up	17	Visual	Apenheul, Twycross	Affiliate/Play
Beckon	3	Visual	Durrell, Twycross	—
Duck lips closed	5	Visual (facial)	Apenheul, Durrell	—
Frog lips	2	Visual (facial)	Apenheul, Twycross	—
Headstand	9	Visual	Durrell, Twycross	—
Hold hand	7	Tactile	Apenheul, Twycross	—
Kiss	4	Tactile	Apenheul, Durrell	—
Look back	6	Visual	Durrell, Twycross	—
Mouth	11	Tactile	Durrell, Twycross	Share (low fidelity)
Play face	5	Visual (facial)	Durrell, Twycross	Affiliate/Play
Present body part	2	Visual	Durrell, Twycross	—
Present genitals	9	Visual	Apenheul, Twycross	—
Pull away	8	Tactile	Apenheul, Twycross	—
Raspberry face	5	Visual (facial)	Apenheul, Twycross	—
Show	4	Visual	Apenheul, Durrell	—
Simultaneous hit	9	Tactile	Apenheul, Twycross	Affiliate/Play
Somersault	42	Visual	Durrell, Twycross	Affiliate/Play
Teeth bared	2	Visual (facial)	Apenheul, Twycross	—
Tense upper lip	2	Visual (facial)	Apenheul, Twycross	—
Whistle face	2	Visual (facial)	Apenheul, Durrell	—

The low frequency of these gestures increases the likelihood that they are present in all three zoo populations, but were not used intentionally during the observation periods of this study. There was not any exchange of individuals between the three zoos studied, so if a gesture exists in only two of the three populations, then it has either developed independently in both or is part of a species-typical repertoire. If all 20 of these "two-zoo" gestures were, in fact, present in all orangutans, then the repertoires of the different populations would be distinguished only by 4 group-specific gestures and 5 idiosyncratic gestures. This would indicate that the gestures of orangutans are overwhelmingly species-typical and that there is little group specificity. Though the repertoires might very similar on the group level, individual repertoires may still vary greatly from animal to animal. The low frequency of many gestures suggests that orangutans' repertoires contain gestures that are only used rarely. Individuals may be able to use or respond to far more gestures than they routinely exhibit. Factors such as age or communicative partner may play roles in determining which gestures are expressed at any given point.

3.7 Age effects

Analysis of the effect of signallers' age on their communicative goals and gestures was performed on the entire set of intentional gestures. Comparison between age groups was similar to the analysis of gesture use between individuals and zoos and used the same dataset. The investigation of age effects covered both specific gestures and signaller goals between different age classes. Importantly, the analysis of signaller goals did not use the measure "observable goal" that was used to determine gesture meanings. Instead, I used "presumed goal" as a measure of the signaller's desires. This more subjective measure was employed for this analysis, because one of my aims was to compare the changing communicative goals across ages. As was mentioned previously, gestures with observable goals are those that appear to be successful. As analysis of the goals of different age classes aims to compare not success, but signaller intent, a measure of "goal" that takes into account how many gestures individuals attempted (rather than how many were successful) must be used. The best measure of intent available was “presumed goal.”

3.7.1 Communicative goals and age of signaller

The social needs of orangutans change throughout their lifetimes. Their communicative desires undoubtedly mirror these changes. Infant orangutans rely on adults (particularly their mothers) for locomotion and food, and spend much of their time playing. Adults, in contrast, spend much of their time responding to infants’ attempts to affiliate or play, asserting their dominance over others, and requesting food or, occasionally, sex. The types of gestures produced by adults and infants almost certainly are affected by the relationship between signaller and recipient as much as by the age of the signaller (i.e. adults request different things from other adults than they do from infants). To illustrate how communicative goals change in response to the relationship between signaller and recipient, the percentage of gestures directed towards different types of goals was calculated for each type of relationship. These frequencies are displayed in Table 19 and Figure 3.

Table 19: FREQUENCY OF GESTURAL GOALS FOR DIFFERENT SOCIAL RELATIONSHIPS.

The percentage of gestures aimed at achieving common goals is presented for each kind of signaller-recipient relationship. The total number of gestures observed for each type of relationship is given in parentheses in the column headings. 1286 gestures were used for this analysis as gestures with unclear goals were excluded.

Type of Goal	Goal	Adult to Adult (115)	Adult to Immature (369)	Immature to Adult (601)	Immature to Immature (201)
Attraction	Play	23%	24%	35%	61%
	Affiliation	10%	8%	17%	15%
	Attention	19%	1%	11%	4%
Directing attention	Look at object	1%	3%	1%	0%
Obtain item	Share food or object	8%	3%	19%	4%
Coordinated Locomotion	Climb on	0%	5%	0%	0%
	Pick up	0%	0%	2%	0%
	Follow	0%	5%	9%	0%
Repulsion	Leave	12%	10%	2%	5%
	Move back	2%	9%	0%	4%
	Stop	10%	32%	4%	7%
Sexual	Solicitation	15%	0%	0%	0%

The different goals fall naturally into larger categories aimed at attracting the recipient, directing the recipient's attention, obtaining something, coordinating locomotion, repulsing the recipient, or initiating mating. The distribution of gestures within these larger categories can be seen through the colour groups in Figure 3.

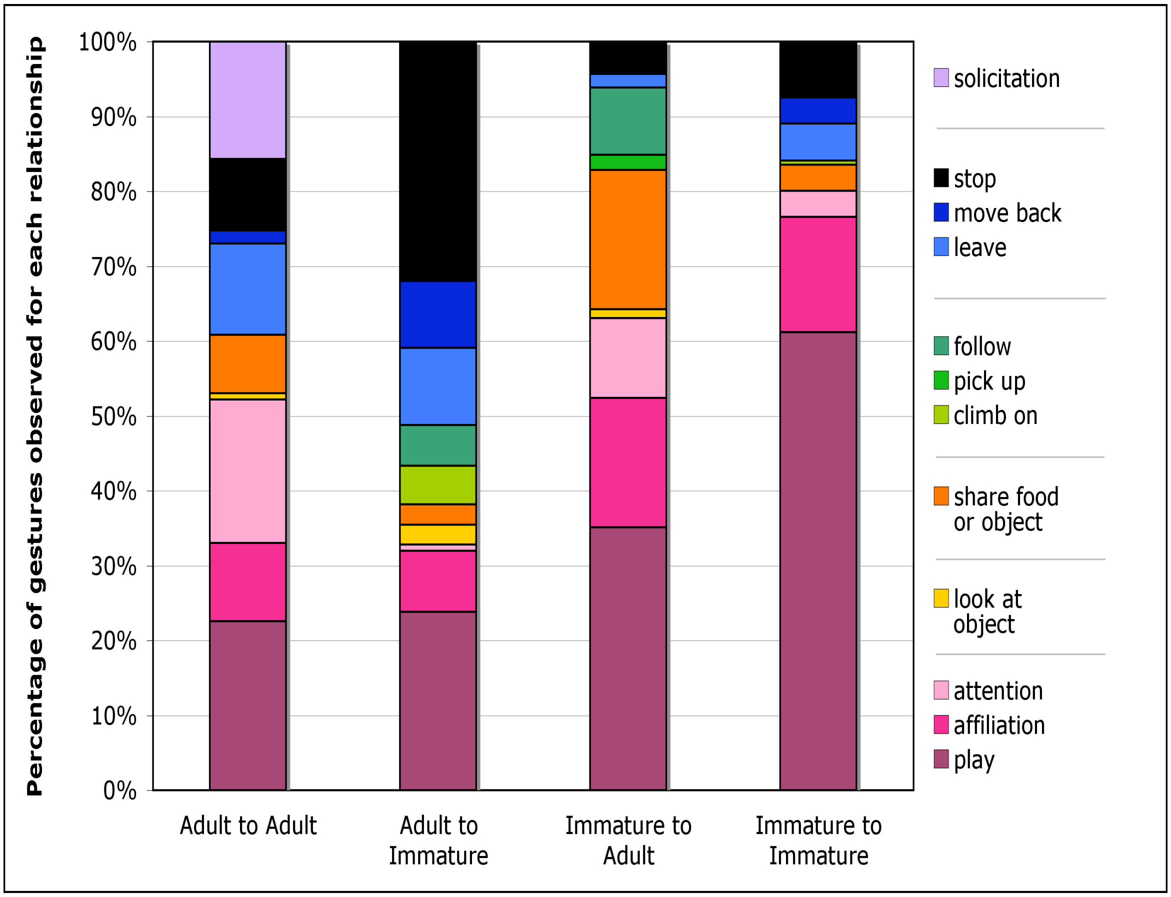


Figure 3: **GOALS OF GESTURES PRODUCED IN INTERACTIONS INVOLVING DIFFERENT TYPES OF SIGNALLER-RECIPIENT RELATIONSHIPS.**

Goals associated with engaging the recipient in positive interactions are displayed in shades of pink, locomotive goals in shades of green, and repulsive and inhibitory goals in shades of blue. As in Table 19, this figure was constructed from 1286 examples of gesture and excludes those with indeterminate presumed goals.

From Table 19 and Figure 3 one can see that immatures communicate mainly to attract attention or elicit positive interactions from either adults or other immatures. The differences in immatures' goals when communicating with adults as opposed to other immatures lie in obtaining items, locomotion, and repulsion. Immatures request items from adults much more often than they do from other immatures. None of the observed gestures between immatures was directed towards locomoting with the other, whereas 11% of the immatures' gestures directed to adults were used to initiate locomotion or carrying. Immatures often attempted to make another immature move away more than twice as often as they tried to move adults away (16% as opposed to 6%). Adults similarly produced gestures in accordance with different goals when communicating with another adult versus an immature. Most strikingly, adults never produced gestures with locomotive goals to other adults, and frequently (32% of the time) produced gestures to immatures with the goal “stop.”

3.7.2 Age-specific gestures

Some actions were only used as intentional gestures by signallers of particular age classes. In certain cases, age-specific gestures may occur because they are directly related to the age-specific goals of individuals in different age classes. In these cases, an individual in an age class that does not typically exhibit a particular gesture might exhibit that gesture if she encountered a situation in which the relevant goal arose. This is likely to be the case with gestures used by infants to solicit food sharing or used by adults to solicit mating. In contrast, age-specific gestures might involve movements used by individuals of all age classes, but are only used as intentional signals by individuals of certain ages. Additionally, some gestures that appear to be limited to a particular age may, in fact, have been observed so rarely that any similarities in their signallers may be purely coincidental. Investigation of age effects is therefore most likely to yield reliable results in the more common gestures found in all three populations.

All non-idiosyncratic gestures that were found to be limited to use by a specific age range (e.g. infants, immatures, adults) are listed in Table 20. Many of the gestures I observed in at least 2 of the 3 zoos were infrequent gestures, and were observed in only a few individuals and fewer than 4 times towards any one goal. In these cases, apparent age effects may be a product of the infrequency of the gesture—if a gesture is used by only two orangutans, those individuals have a fair chance of belonging to the same age group even if chosen randomly.

Table 20: GESTURES OBSERVED ONLY IN CERTAIN AGE CLASSES.

The gestures, their meanings (if they met the threshold for consistently predicting an observable goal) are included in the table along with the age class of the signaller. A value of "—" in the observable goal column indicates that the gesture was observed fewer that 4 times in conjunction with all goals and was thus excluded from analysis of meaning. In total, the species-typical gestures in the table were calculated from 126 tokens, and the gestures found in 2 zoos were calculated from 120 tokens.

	Gesture	Meaning	Age Class
"Species-typical" (found in 3 zoos)	*Poke*	—	Infants/juveniles
	Air-bite	Affiliation/play	Immatures
	Cover	—	
	Food-beg-orally	Share	
	Roll-on-back	Affiliate/play	
	Embrace-pull	Climb on	Adult
	Back roll	Affiliate/play	Non-infants
	Nudge	Move back/leave	
	Put object on head	Affiliate/play	
	Shoo	Move back/leave	
Found in 2 zoos	*Arms up*	Affiliate/play	Immatures
	Present genitals	—	
	Show	—	
	Somersault	Affiliate/play	
	Duck lips closed	—	Adults
	Look back	—	
	Whistle face	—	
	Frog lips	—	Non-infants
	Present body part	—	
	Raspberry face	Affiliate/play	
	Simultaneous hit	Affiliate/play	
	Teeth bared	—	
	Tense upper lip	—	

Gestures found in all three zoos

Ten of the 35 gestures observed in all three zoos were used only by individuals of a certain developmental stage. *Poke* was used only by infants and juveniles (animals under 6 years old). *Air-bite*, *cover*, *food-beg-orally*, and *roll-on-back* were used only by immatures (infants, juveniles, and sub-adults under 12). *Embrace-pull* was used only by adult females to their offspring. The gestures *back-roll*, *nudge*, *put-object-on-head*, and *shoo* were never observed to be used intentionally by infants.

The rarity of some of these gestures may mean that the observed age effects are a product of sampling bias (particularly in the case of the gestures observed only in infants and juveniles), so interpretation of the distribution of rare gestures must be tentative. However, the frequent gestures found to be limited to a specific age range seem to be a true reflection of the difference in social goals between individuals of different ages. In the case of the adult-specific gesture, *embrace-pull*, the gesture appears to be particularly useful for guiding young orangutans towards the signaller while providing them an opportunity to climb onto the signaller as well. If providing an opportunity for the recipient to climb on is an important part of the goal, then the use of this gesture may be determined by the age of the recipient as much by the age of the signaller.

The fact that *food-beg-orally* was used only by immatures should not be surprising as food-begging would be expected to decrease with age, and oral begging is most often seen in infants.

The gestures *nudge* and *shoo* are used to displace other individuals or to make them move back. It is likely that these gestures were not observed in infants because infants do not often try to move others away, but more often seek to interact with them. Two other gestures absent in infants are *back-roll* and *put-object-on-head*. Infants were observed to use these movements during play, but never as directed gestures with the anticipation of a response. It is possible that over time, orangutans learn to use these play movements as communicative signals as other individuals respond to them as an intention to play. Perhaps infants have not had enough experience using these movements with other individuals to have established them as signals in addition to movements performed during solitary play.

Group-specific gestures

None of the gestures identified as group-specific were used only by individuals of a particular age class or limited to situations involving the same relationship between signaller and recipient.

Gestures found in two zoos

Gestures that were documented in two of the three zoos were often observed only in individuals of a specific age class. However, most of the time, these gestures were observed only in one individual in each of the two collections and the similarity of age class may be coincidental. To lower the probability of similar age due to chance, I required that a gesture be observed in at least 3 individuals of similar age (two in one zoo and one in the other) in order to be counted as an "age-specific" gesture. Four gestures that were found in two zoos fit met this criterion: *arms up, somersault, raspberry face,* and *simultaneous hit*.

The gestures a*rms up* and *somersault* were both only observed in immatures (infants, juveniles, or subadults), while *raspberry face* and *simultaneous hit* were never observed in infants. All 4 gestures were used to request affiliation or play. Adults were observed using *somersault* as an action, but never as an intentional gesture. The two gestures not observed in infants, *raspberry face* and *simultaneous hit*, may require practice or involve motor skills the infants have not yet developed. I am not aware of whether infants are able to make the raspberry vocalization. I never observed an infant producing the noise, so perhaps it takes practice or advanced muscular control to develop. If the *raspberry face* gesture is a silent use of the vocalisation, then orangutans might not use the facial expression as a communicative gesture until after they have mastered the vocalisation. The gesture *simultaneous hit* may require practice as well as it involves timing one's movements with another individual. It is possible that infants have not yet mastered this ability and so are not able to produce *hit* simultaneously with a partner.

3.8 Discussion

Once non-intentional gestures had been excluded and additional structural and social variables were used to help define new gestures, the orangutans in our study were found collectively to use 25 tactile gestures, and 39 visual gestures (including 12 facial expressions). One projectile gesture (*throw*) was observed, but as it was only observed a single time in a situation that met our criteria for intentional usage, it was discarded from the analysis. The dataset of intentional gestures was created by using our criteria for intentionality to exclude potential gestures that were not directed towards another individual, goal-directed, and flexible (used strategically with different gestures or in response to the recipient's behaviour). By employing these criteria, 9 potential gestures were determined to be non-intentional and excluded from analysis. Several datasets were used throughout the thesis for different analyses, a summary of which analyses were performed using which data is provided in Appendix V.

The measure "observable goal" was created to combine a subjective variable that measured signaller's motivation ("presumed goal") with an objective variable that measured recipient's response ("outcome"). This new variable was intended to identify gestures that were successful so that their meanings could be analysed in depth. This analysis identified which gestures appeared to have specific meanings and also which variables were important in determining the meaning of a gesture. The comparison of social and structural variables provided insight into the cognitive process of both signaller and recipient in using and interpreting orangutan gestures. Through the comparison, it was possible to identify cases where one variable was the most effective in predicting meaning (as in the case of *seize*) and to determine the importance of social and contextual variables in changing the meanings of gestures. To communicate effectively, orangutan signallers must monitor the recipient's behaviour, and visual attention, and choose appropriate signals. Likewise, orangutan recipients must focus on not only the form of the signal, but also the identity of the signaller.

The creation of the “effectiveness index” allowed me to directly compare the influence each variable had over the meaning of the gestures. This allowed me to identify “target” as the most important structural variable and “relationship between signaller and recipient” as the most important social variable in predicting the meaning of gestures. The ability to rank potentially significant variables such as “target” and “relationship” allows us to generate hypotheses about the boundaries between gestures and would make it possible to quantify similarity or dissimilarity between gestures based on the most salient variables.

3.8.1 Further analyses: using gesture to understand cognition

All of the gestures analysed in this chapter were treated as individual gestures in order to determine meaning and investigate the effects of group membership or age on repertoire. However, of the 1421 intentional gestures we observed, 940 (66%) appeared in sequences of at least 2 gestures. Sequences of gestures might be used by orangutans to communicate more complex desires, or to clarify and/or amplify their meaning when they have been unsuccessful in eliciting a desired behaviour from the recipient. Clarity, rather than complexity, seems a more likely driving force behind the construction of sequences of gesture. But only a comparison of homogeneity of meaning within sequences would allow one to conclude whether sequences represent persistence towards a goal or communication of a more complex goal.

It is apparent that orangutans use gestures in combination with one another, but not clear under what circumstances sequences arise. Of the original 1334 intentional gestures, 506 (38%) received no immediate response, and an additional 107 (8%) caused the recipient to resist or turn away. Studying the sequences of gestures signallers produce will help to shed light not only on the motivation behind the use of gestural sequences, but on what social information orangutans take into account when choosing subsequent gestures. Investigation of the use of different gestural modalities in sequences would make it possible to determine whether orangutans extend their awareness of the signaller's visual attention into the recent past and choose gestures based on what recipients have already seen. Similarly, comparison of sequences produced towards unresponsive recipients versus recipients who actively reject or avoid the signaller's communicative attempts would indicate whether orangutans employ different strategies towards unaware versus unwilling recipients. Identification of such strategies in the persistence and elaboration of initial gestures would provide great insight into the mind of the signaller and shed light on the relationship between communication and cognition.

Chapter 4: Use of gestural sequences

4.1 What can sequences of gestures tell us about cognition?

The analysis of individual gestures can reveal much about the meaning of specific signals, but does not take into account the dynamic environment in which gestures occur. It is difficult to consider the nature of a communicative act out with the context in which it was produced (most often an ongoing exchange between two individuals). In fact, it has been suggested that it is the interaction between individuals, rather than the behaviour of one party, that creates the meaning of particular actions (Griffin 1997; King 2004). In order to understand how gestures fit into interactions between orangutans, it is necessary to examine them as parts of the contexts in which they occur, taking into account the recipient's as well as the signaller's actions. Just as the identity of the recipient and the recipient's behaviour prior to the gesture are influential in determining the form or success of a gesture, the recipient's actions or reactions during a communicative exchange likely influence how the signaller gestures. In addition to responding to the behaviour of the other individual, orangutan signallers might also be influenced by the nature and success of their own communicative attempts, remembering what gestures they tried and how they were received. If an orangutan produces more than one gesture to another orangutan, it becomes possible to examine not only which gestures were used, but also whether the gestures were employed in predictable combinations and whether they were used to reiterate meaning or change meaning.

In evaluating signals of non-verbal individuals (i.e. infants and animals), a persistent desire to communicate a goal and a diversity of methods employed in reaching it are the best indicators of intentional communication we have (Bates et al. 1979; Golinkoff 1986; Lock 2001; Leavens et al. 2005). Sequences of communicative signals thus can provide better evidence of intentional communication than signals produced on their own. In addition to evidence of intentionality, sequences of gestures produced when a goal remains unfulfilled may allow us to evaluate the signallers' communicative strategies, such as what they understand about the recipient's state of mind.

To date, non-human gestural sequences have only been investigated in chimpanzees. Two studies on the overall use of gesture by chimpanzees determined that chimpanzees often precede visual gestures with auditory or projectile gestures in order to attract the attention of the recipient who is not looking towards them (Tomasello et al. 1994; Leavens et al. 2004). However, a study of chimpanzee 2-gesture combinations found no evidence that chimpanzees used certain gestures to attract the attention of others before using visual gestures (Liebal et al. 2004). In all of these studies, the authors assumed that gestures used sequentially were directed towards the same goal and communicated the same meaning, yet no evidence was ever presented to support this viewpoint. In an attempt to be objective, the analyses presented in this chapter were begun without any assumptions as to the functions of sequences. As in the analysis of individual gestures, I aimed to let the behaviour of the orangutans inform me about the meaning and use of sequences.

In order to determine what sequences of gestures indicate about the goals and decisions of the signaller, it is necessary to first investigate the circumstances under which sequences are produced—whether they are indicative of a signaller's unfulfilled goals, or perhaps could be used to communicate different meanings than individual gestures. If orangutans use multiple gestures when the initial goal of the signaller has not been met, it would indicate that multiple gestures arise as a function of the persistent desire to achieve a specific goal. However, it is also possible that gestures are combined into sequences in order to convey a meaning that is different from that of any single gesture. If orangutans use multiple gestures to convey a persistent desire for a goal that has not yet been fulfilled, sequences should arise primarily when the recipient does not respond or responds in a way that does not fulfil the goal of the first gesture. Also, one would expect that gestures later in a sequence would have the same meaning as the first gesture. If, on the other hand, sequences are used to convey meanings beyond that of the first gesture, then the response of the recipient to the first gesture should not affect the likelihood of persistence and the gestures in the sequence would not necessarily have the same meaning as one another. If sequences of gestures were found to have different meanings from the gestures themselves, it would imply that orangutans have the ability to combine gestures according to some syntactic rule. The ability to combine meaningful communicative signals to create meanings distinct from those of the original signals has only been found in one non-human primate species (a monkey) (Arnold and Zuberbuhler 2006), and it is often speculated that certain elements of syntax represent the greatest difference between human language and non-human primate communication (Kako 1999; Hauser et al. 2002). Given the lack of meaningful (as opposed to phonological) syntactic elements in vocal communication systems of non-human primates, it is a priori unlikely that orangutans use sequences of gestures to create new meanings. However, to rule out this possibility, it is necessary to investigate under what circumstances multiple gestures are used, and whether gestures within a sequence most often have the same meaning as one another.

Orangutans might merely repeat their gestures if they do not achieve the desired response, or they may persist strategically, employing different types of persistence in different contexts. If orangutans use multiple gestures in situations where their initial goals have not been fulfilled, then when and how they persist may reveal what variables they take into account when choosing how to communicate. Signallers may choose to continue signalling in all cases where the recipient fails to meet their goals, or only in those cases where the recipient fails to respond at all. Provided sequences of gestures are used primarily in a continuing attempt to achieve an unaltered objective, the strategies signallers use in different social situations may provide insight into the orangutan's ability to communicate in response to the behaviour, gaze, or knowledge states of others.

4.2 Methodology

To investigate the function of sequences of gestures produced by orangutans, I analysed the sequences of gestures that individuals made during dyadic interactions. ***Gesture sequences*** were defined as the signaller's gestures directed towards the same recipient without any pauses longer than 10 seconds between gestures. A sequence was considered to have ended when more than 10 seconds elapsed between gestures, or when the signaller turned away from the recipient or left.

SEQUENCE: Series of gestures produced by one individual and directed towards the same recipient. Ends when signaller pauses for more than 10 seconds or turns away.

4.2.1 Variables analysed

Social and structural variables were coded for each gesture in order to contextualize it within a sequence and exchange between two individuals. The variables used for this analysis (also listed in Table 3) were:

<u>Recipient's actions</u>

1. **Response**
2. **Outcome of interaction** (final recipient response)

Signaller's actions following initial gesture

1. **Response waiting** (signaller pauses for >2 seconds after gesture while remaining oriented towards recipient)
2. **Use of another gesture** (signaller uses another gesture toward recipient within 10 sec)
3. **Relationship of next gesture to current gesture** (whether the next gesture is a form of repetition of the current gesture, is a different gesture of the same modality, or is a gesture of a different modality)
4. **Other gestures used** (each within 10 seconds of the previous one)

Context of gesture within exchange

1. **Number of gestures in sequence**
2. **Position of gesture in sequence**
3. **Role of the signaller in communicative exchange** (e.g. initiator)

The recipient's ***response*** was defined as the most salient change in the recipient's behaviour in the 10 seconds following the signaller's gesture. If a more active response followed a subtler one within the 10 seconds, it was considered to be the response. Thus if a recipient first looked towards and then *hit* the signaller following a gesture, *hit* (a gesture) was considered to be the response. If the recipient gestured multiple times in response, the recipient's first gesture was deemed to be the response. If the signaller produced a second gesture before a noticeable change in the recipient's behaviour, the first gesture was recorded as receiving no response (even though the recipient may have reacted within 10 seconds). In contrast, ***outcome*** was defined as the net result of an interaction. Thus, if several gestures were employed in a sequence and the recipient responded by playing with the signaller after the 3rd gesture, the outcome of all three gestures in the sequence would be "play," but "play" would only be the response for the third gesture. The measure of outcome was also used in determining "observable goals" for gestures for the analysis of meaning in Chapter 3.

RESPONSE: Most active change in the recipient's behaviour within 10 seconds of a gesture but before the signaller gestures again.
OUTCOME: End result of interaction.

In contrast to the analyses in Chapter 3 that involved only those gestures used intentionally, the analyses of sequences were performed on the natural sequences of acts, regardless of whether they had demonstrated clear intentional usage. The dataset used for these analyses included a total of 1421 gestures (87 more than the 1334 purely intentional gestures). This was done so that the signallers' strategies could be investigated as they had occurred rather than excluding actions that had failed to meet the criteria for intentionality from the middle of sequences thereby creating unnatural patterns. A summary of the different datasets of gestures and the analyses that were performed using them is given in Appendix V.

4.3 Why use more than one gesture?

Orangutans often use multiple gestures when communicating with one another. In this study, orangutans produced 730 gestural "utterances" (solitary gestures or sequences).

GESTURAL UTTERANCE: Gestures produced by one individual directed towards another within 10 seconds of one another during the same interaction. May be either sequences or solitary gestures.

Of the observed utterances, 349 were sequences of at least 2 gestures. These 349 sequences contained 1042 incidents of gesture and ranged in length from 2 to 16 elements. The average length of utterances was 1.73 gestures (± 1.23). The frequency of all utterance lengths is displayed in Figure 4.

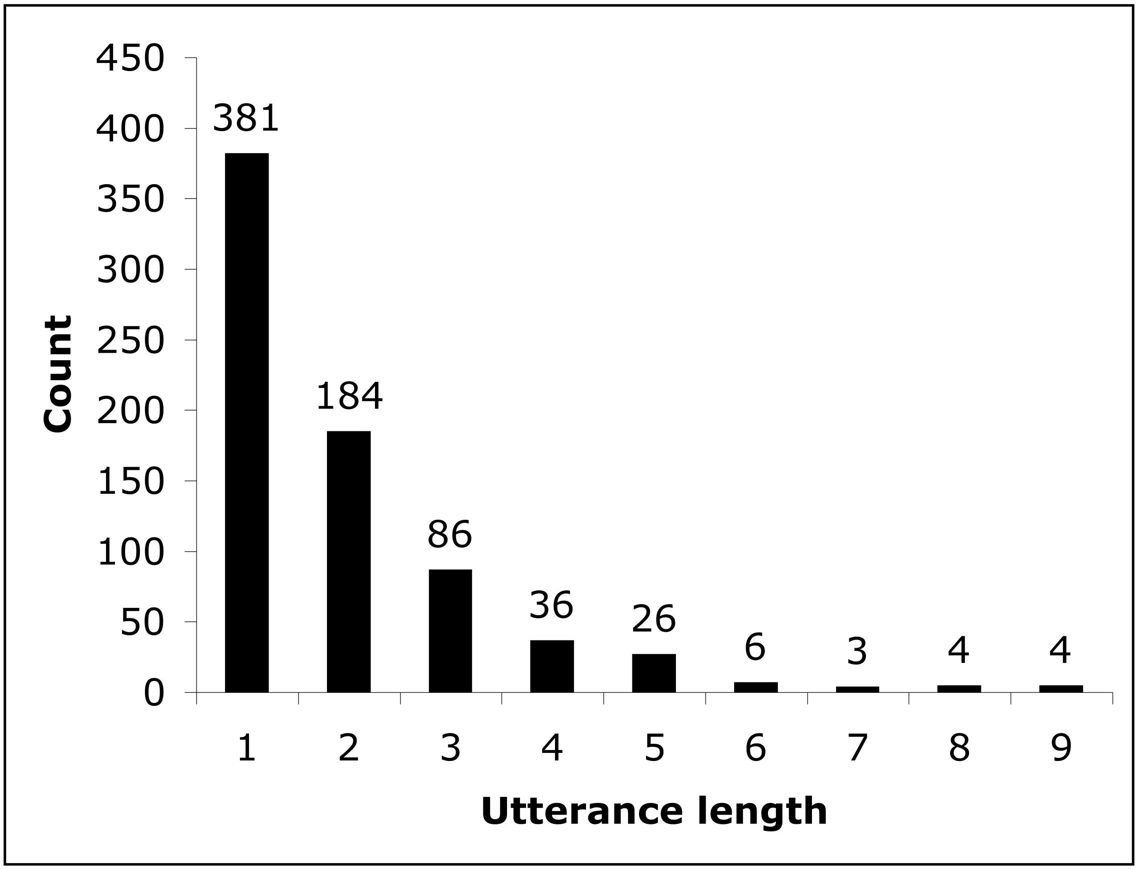


Figure 4: LENGTH OF ALL GESTURAL UTTERANCES.

The most frequently observed utterance length was a single gesture. It is possible that the solitary gestures all represent interactions where the recipient fulfilled the signaller's goal and the multiple gestures indicate exchanges where the signaller continued gesturing until her goal was fulfilled or she gave up. It is also possible, however, that solitary gestures were used when the signaller had a simple, easily-transmittable goal, and multiple gestures were used to communicate different or more complex meanings. Utterances may also have been recorded as solitary gestures when a signaller paused for more than 10 seconds between gestures, so some solitary gestures were likely used during exchanges in conjunction with other gestures but did not meet the criteria for inclusion in a sequence.

4.3.1 Do sequences reinforce or alter meaning?

In order to determine whether gestural sequences were strategic choices made when an initial gesture failed to receive the desired response, or were combinations of gestures that changed the overall meaning of the gestural utterance, I examined the consistency of meaning between the first and second gestures produced. Sequences in which the first and second gestures have the same meaning indicate that the signaller is probably persisting in an attempt to reach an initial goal rather than attempting to create a new meaning. If gestural sequences are an indication of persistence towards a goal, a high frequency of redundancy of meaning in the gestures should provide support for the analysis of meaning of specific gestures used in this study (see Chapter 3).

I examined the meanings of the first and second gestures in sequences at least 2 gestures long (Table 21). I investigated the frequency of only those gestures determined to have predictable meanings (Table 8). The frequency of gestures with no consistent goal was noted, but as it was unclear what meanings (if any) these gestures had, they could not be used to determine whether the meaning of the second gesture matched that of the first. To analyse consistency of meaning in sequences, I ascribed meaning to each of the 29 high-fidelity gestures and the 7 low-fidelity gestures, which were identified in Chapter 3 by their consistent correlation with one observable goal (Table 8). For the purposes of this analysis, high and low-fidelity gestures were combined into one category of gestures that "predictably" had a single meaning. The gestures that failed to meet the criteria for prediction of a goal were recorded as having "no consistent goal." It was assumed that these gestures were too infrequent to meet the criteria for predicting a goal, or were used frequently in more than one context.

Sequences beginning with the two most common meanings—"affiliate/play" and "move back/leave"—were analysed to determine whether the second gesture used had the same meaning or a different meaning. The frequency of meanings in second gestures was analysed for each individual for sequences beginning with both "affiliate/play" and "move back/leave." The average number of sequences beginning with "affiliate/play" for each individual was 15.19 (± 17.55), and for “move back/leave” was 2.71 (± 2.05). The overall frequencies (across individuals) of these two meanings as first gestures in sequences were 73% for “affiliate/play,” and 17% for “move back/leave.” The frequency of second gestures having the same meaning was analysed for individuals who exhibited at least 5 sequences beginning with “affiliate/play” and for those who exhibited at least 2 sequences beginning with “move back/leave.” This narrowing of the dataset was performed to restrict analysis to only those individuals who had performed enough sequences to generate reliable frequencies of second meanings. As the average number of sequences per individual was only 2.71 for sequences beginning with “move back/leave,” the lower cut-off of more than one sequence was used in an attempt to avoid over- or underestimating proportions through lack of data. Expected frequencies for second gestures of both meanings were the same as their observed frequencies as first gestures. Sequences in which the second gesture was a repetition of the first were excluded from the analysis as meaning was, by definition, the same but could have been the result of excitement rather than indicating a choice of an equivalent gesture. A breakdown of the proportion of second gestures persisting in those meanings and the comparison to the expected frequencies is presented in Table 21.

Once repetitive sequences and individuals exhibiting fewer than 5 sequences beginning with “affiliate/play” or 2 sequences beginning with “move back/leave” were excluded, there were 203 sequences beginning with gestures meaning “affiliate/play,” 184 of which used a second gesture of the same meaning. There were 14 sequences beginning with gestures meaning “move back/leave,” 9 of which had second gestures of the same meaning.

Table 21: Meanings of first and second gestures for each individual in sequences beginning with gestures meaning "affiliate/play" or "move back/leave"

Each row describes the sequences produced by one individual. The percentage of second gestures with the same meaning as the first is calculated for each individual out of the total number of second gestures that were deemed to have a persistent meaning. Gestures with unclear or ambiguous meanings were not included in the calculation of frequency. The expected frequencies in the right hand column were taken from the proportion of meaningful first gestures that had the meaning "affiliate/play" (73%) or "move back/leave" (17%).

First Gesture	**Second Gesture**				
	Same	Different	*Unclear (ambiguous gestures)*	% same meaning (out of meaningful gestures)	Greater than expected (73%)/(17%)
Affiliate/Play	2	0	*3*	100%	Y
	3	1	*1*	75%	Y
	3	0	*1*	100%	Y
	5	0	*0*	100%	Y
	8	0	*4*	100%	Y
	13	2	*13*	87%	Y
	6	1	*4*	86%	Y
	13	0	*8*	100%	Y
	3	0	*3*	100%	Y
	6	1	*1*	86%	Y
	5	2	*4*	71%	
	21	0	*5*	100%	Y
	18	3	*10*	86%	Y
	5	1	*2*	83%	Y
	2	1	*3*	67%	
	5	0	*1*	100%	Y
	4	0	*2*	100%	Y
	17	3	*18*	85%	Y
	45	2	*15*	96%	Y
Move back/leave	1	0	*1*	100%	Y
	0	1	*2*	0%	
	3	0	*0*	100%	Y
	2	0	*0*	100%	Y
	1	1	*1*	50%	Y
	1	1	*1*	50%	Y
	1	1	*3*	50%	Y

When the first gesture in a sequence was a gesture of the type “affiliate/play,” the average probability of the second gesture’s sharing this meaning was 91% (SD 11%). The second gesture had a different meaning only 9% of the time (SD 11%). For sequences beginning with “move back/leave” gestures, the second gesture had the same meaning 64% of the time (SD 38%), and a different meaning 36% of the time (SD 38%). Figure 5 compares the expected frequency of gestures meaning “affiliate/play” (73%) to the observed frequency (for each individual) of sequences starting with “affiliate/play” gestures containing second gestures with the same meaning. Figure 6 compares the expected frequency of gestures meaning “move back/leave” (17%) to the observed frequency (for each individual) of sequences starting with gestures meaning “move back/leave” containing second gestures with the same meaning. It is clear from Table 21, that second gestures (in sequences beginning with either meaning) were also frequently of a type that did not have a consistent meaning. Since gestures with low success rates, rare gestures, and ambiguous gestures were all included in the category of gestures with unclear meanings, it is impossible to determine whether cases in which the second gestures had unclear meanings are the result of meaningful gestures followed by gestures with different or multiple meanings, or followed by gestures the same meaning but which failed to meet one of the thresholds for the analysis of meaning.

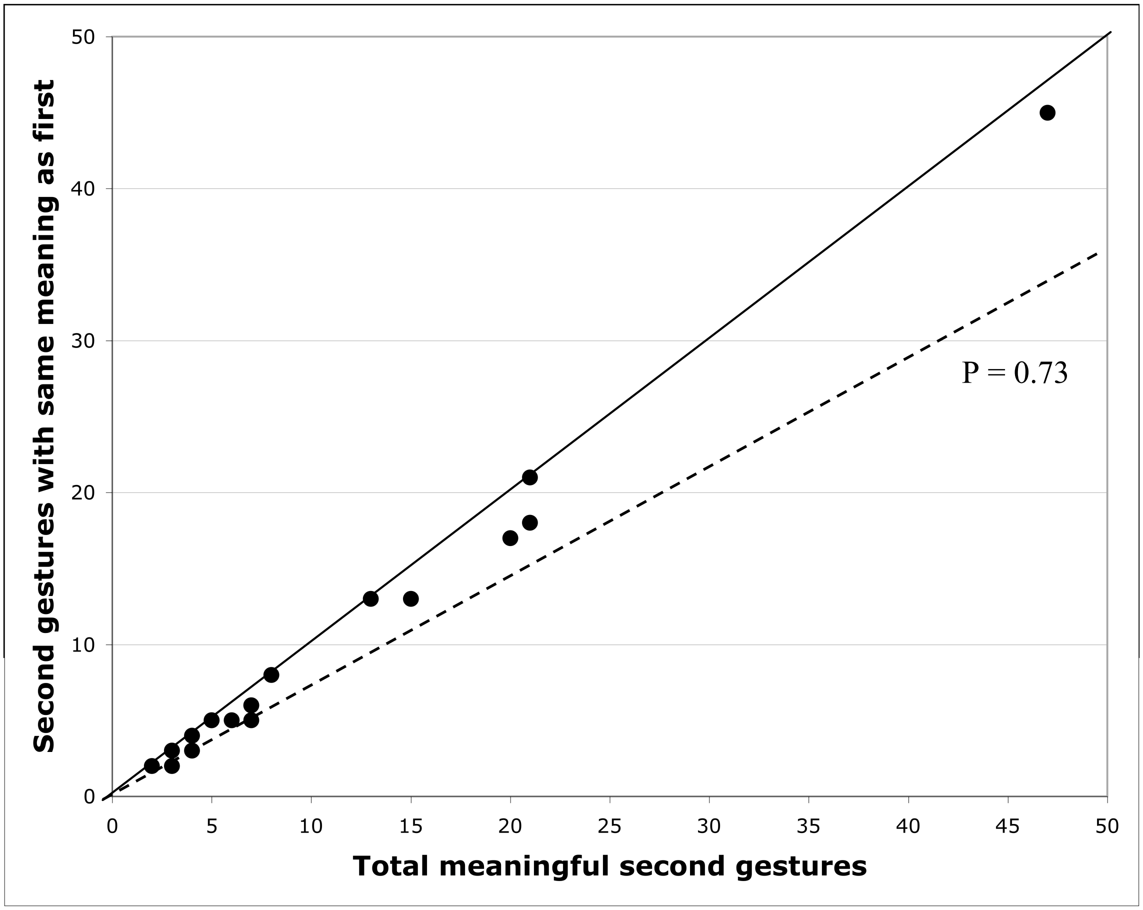


Figure 5: FREQUENCY OF SECOND GESTURES DUPLICATING MEANING OF THE FIRST FOR SEQUENCES BEGINNING WITH "AFFILIATE/PLAY."

Each point indicates the gestural sequences produced by one individual. The dashed line represents the probability of a meaningful gesture having the meaning "affiliate/play" based on the frequency of first gesture meanings (0.73). The solid line represents perfect concordance between the meanings of the first and second gesture, i.e. the distribution expected if the meaning of the second gesture were always the same as that of the first. Points that lie above the dashed line have an observed frequency greater than that predicted by the expected frequency of gestures meaning "affiliate/play," those below it have an observed frequency less than the expected value.

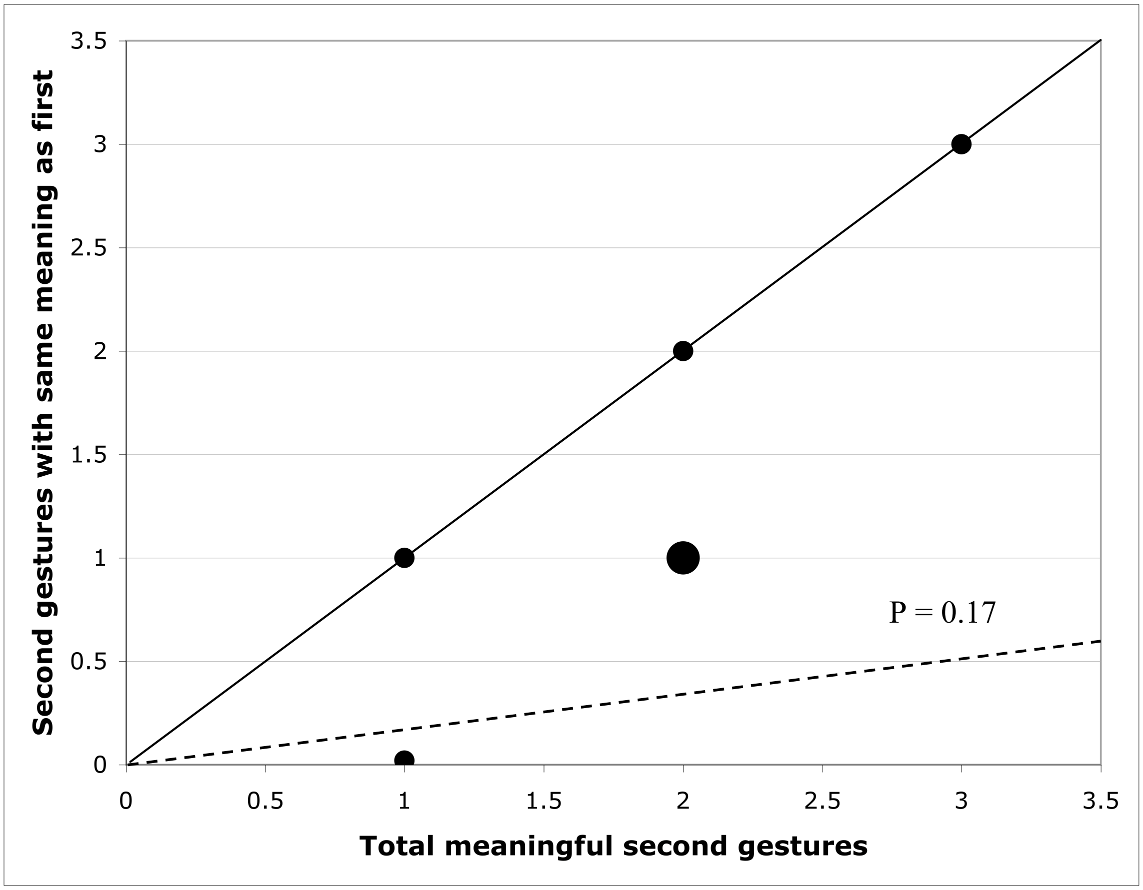


Figure 6: FREQUENCY OF SECOND GESTURES DUPLICATING THE MEANING OF THE FIRST FOR SEQUENCES BEGINNING WITH "MOVE BACK/LEAVE."

Each point indicates the gestural sequences produced by one individual. The dashed line represents the probability of a meaningful gesture having the meaning "move back/leave" based on the frequency of first gesture meanings (0.17). The solid line represents perfect concordance between the meanings of the first and second gesture, i.e. the distribution expected if the meaning of the second gesture were always the same as that of the first. The larger point represents 3 individuals with equivalent values. Points that lie above the dashed line have an observed frequency greater than that predicted by the expected frequency of gestures meaning "move back/leave," those below it have an observed frequency less than the expected value.

Since the majority of individuals used second gestures of the same meaning as the first more often than chance (i.e. their points lay above the dashed lines on Figures 5 and 6), data from all individuals was pooled for analysis of frequency. The mean frequencies of initial meaningful gestures having either "affiliate/play" or "move back/leave" as meanings were used as the probabilities that second gestures would exhibit those meanings (.73 and .17 respectively). Binomial tests were used to compare the observed frequency of same-meaning second gestures to the number expected by their frequency as first gestures. Second gestures matched the meanings of the first gestures significantly often both for sequences beginning with gestures meaning “affiliate/play” (Binomial test, n=203, p<0.0001) and sequences beginning with gestures meaning “move back/leave” (Binomial test, n=14, p=0.0001).

The second gesture in a sequence had the same meaning as the first significantly more often than was predicted by the frequency of both meanings as first gestures. This suggests that when orangutans use multiple gestures, they normally do so in an attempt to reach their initial goal. Since the meanings of first and second gestures in a sequence were more likely to reinforce than to contradict one another, it is reasonable to conclude that when gestures are used in a sequence it is usually with the intention of clarifying or emphasizing the original goal rather than combining gestures to create new meanings.

4.3.2 Does persistence pay?

When orangutans use more than one gesture in a sequence, their second gestures are most likely to have the same meaning as the first. This persistence in meaning is likely indicative of a persistent goal—one that was not fulfilled after the first gesture. However, if a recipient fails to meet the signaller’s goal after the first gesture, will subsequent gestures increase the probability of its being met? To put it another way: if an orangutan attempts a gesture and fails to receive the desired response, can any benefit be expected to result from trying again? If signallers are not likely to achieve their goals by continuing to gesture, then orangutans’ use of different strategies in their persistence might be by-products of other cognitive processes and hold little communicative benefit in themselves.

In order to determine whether there was any benefit to producing sequences, all gestural tokens that occurred within sequences were coded for both the recipient's immediate response and for the recipient's final response to the entire interaction or sequence (i.e. outcome). This measure of outcome indicates whether the signaller eventually received any reaction from the recipient and serves a measure of the net outcome of an interaction. Since each incidence of a gesture was coded for both an immediate response and an eventual outcome, it was possible to examine gestures in sequences and compare the likelihood of receiving a certain type of response immediately after a gesture to the likelihood of receiving it eventually (Figure 7).

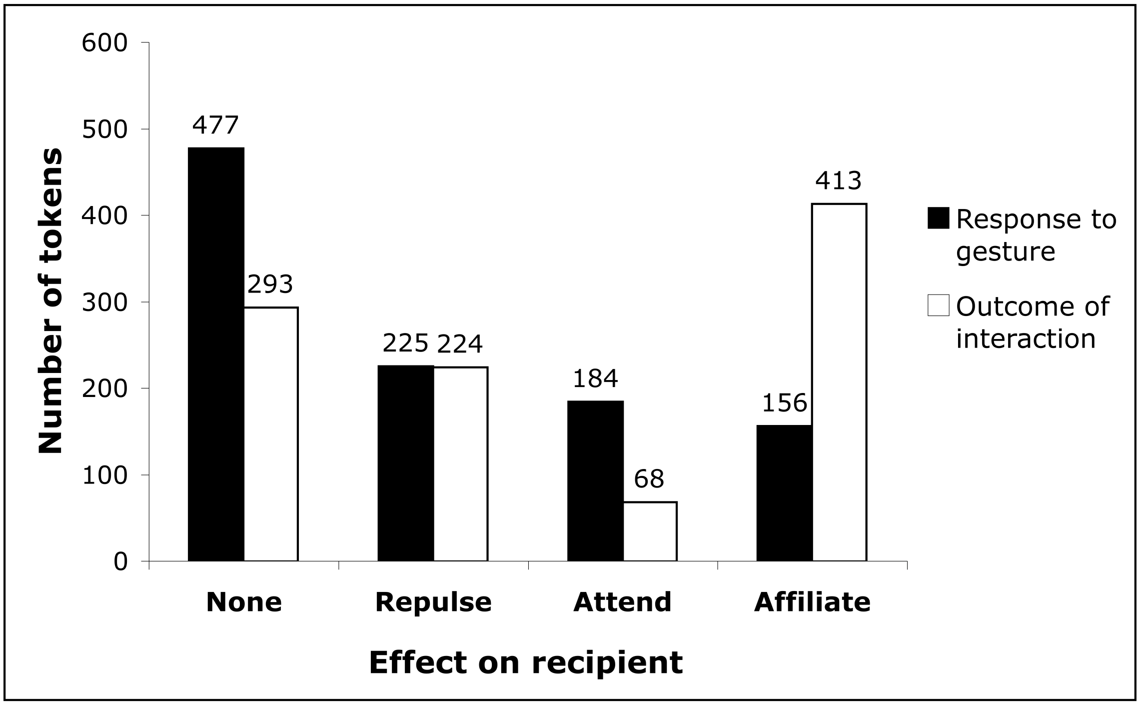


Figure 7: FREQUENCY AND NATURE OF *RESPONSE* VERSUS *OUTCOME*.

One type of outcome, "stop behaviour," was not included in this figure since it did not reliably fall into any of the categories of response. There were 44 observations of this behaviour. Because of this, the displayed responses outnumbered the outcomes (1042 to 998).

Chi-squared tests were used to compare the relative frequencies of eliciting different behaviours from the recipients as immediate responses versus eventual outcomes. Three of the four types of behaviour varied according to whether the behaviour was an immediate response to a gesture or the eventual outcome of a sequence. Recipients were more likely affiliate with the signaller as the outcome of a sequence than as an immediate response to any particular gesture ($\chi 2=176.8$, df=1, $p<0.005$). Conversely, recipients were more likely to immediately respond to a gesture by looking towards the signaller ("attend") than they were to have attention be the eventual outcome of the interaction ($\chi 2=55.4$, df=1, $p<0.005$). Signallers were also more likely to receive no reaction from the recipient as an immediate response than they were as an eventual outcome ($\chi 2=58.5$, df=1, $p<0.005$). The frequency of signallers repulsing recipients (causing them to look or move away) did not change as an effect of response versus outcome ($\chi 2=0.023$, df=1). All probabilities were Bonferroni-corrected for multiple testing.

These findings demonstrate that by persisting in a sequence, a signaller raises her chance of obtaining an affiliative response eventually and lowers her chance of receiving no response. The increase in likelihood of an eventual positive response illustrates the success of persistence. If an orangutan does not achieve her goal on one attempt, she is likely to gain a positive response eventually by persisting in her attempts.

4.4 Do orangutans take the recipient's actions into account when choosing to persist?

Orangutans are more likely to receive a positive reaction from the recipient as a final response to a sequence than as the immediate response to any one gesture, so there is some evidence that continuing to communicate is effective at achieving a desired response. It is unclear, however, whether the recipient's actions play any role in determining whether the signaller will persist or not. When orangutans perform additional gestures in an attempt to reach a communicative goal, they are not broadcasting signals non-directionally into space, but rather to other individuals who may respond in very different ways. It is possible that these responses (or lack of response) might influence the ways in which orangutan signallers persist in their communication.

4.4.1 Are orangutans more likely to persist when the recipient does not respond?

Since sequences of gestures are most often used to forward a single communicative goal, they should occur most often in cases where the recipient's goal was not fulfilled on the first gesture. Exchanges in which the recipient does not respond in any way to the first gesture are instances where one can confidently conclude that the signaller's goal was not met. If the recipient does not respond in any way, the signaller's goal cannot have been fulfilled, so the frequency of persisting to communicate should be higher. To determine whether there was an influence of the recipient's lack of response on orangutans' chance of using multiple gestures, I compared the probabilities of the signaller's attempting another gesture in cases where the recipient did and did not respond. By using the probabilities of the signaller's adding another gesture at different points in the sequence (gestures 1-4), the differences in the likelihood of persisting are easily visible (Figure 8).

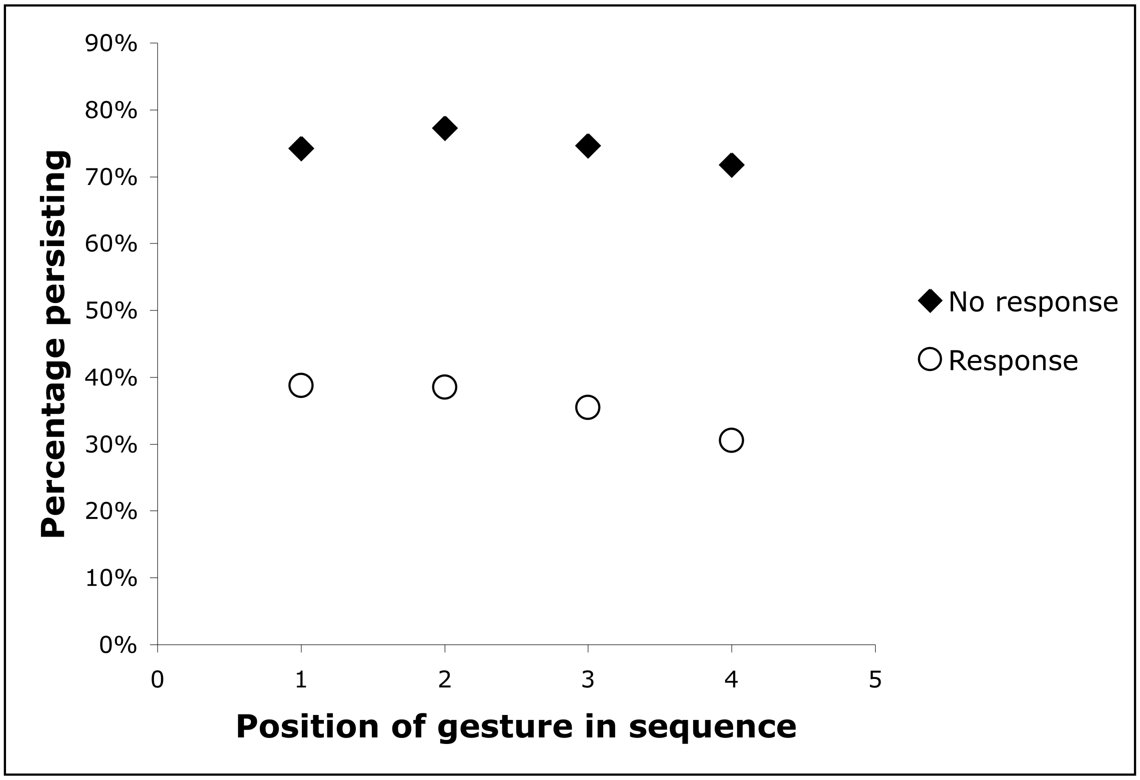


Figure 8: PROBABILITY OF CONTINUING TO GESTURE BASED ON RECIPIENT'S RESPONSE.

The number of observed gestures decreases with every position in the sequence because there are fewer sequences of longer length than there are solitary gestures. There were 730 first gestures and only 82 fourth gestures.

Only gestures in positions 1-4 in a sequence were examined. There were only 45 sequences that were at least 5 gestures long, compared to 349 sequences at least 2 gestures long, 168 at least 3 gestures long, and 82 at least 4 gestures long. The chance of using another gesture following no response from the recipient ranged from 72% to 77% (mean 74% ± 2%). Following any type of recipient response, the chance of persisting ranged from 31% to 39% (mean 36% ± 4%). A paired t-test was used to compare the probability of persisting at the four first positions in a sequence following either a lack of response or any response. A significant difference between the rates of persistence was found between the two conditions (Paired t-test, df=4, t Stat=8.6, p=0.001). The probability of an orangutan's continuing to gesture was significantly greater following no response from the recipient than it was following some response. This analysis and

Figure 8 takes into account only the number of gestures that have already been attempted, not the recipient's responses to past gestures; however, there did not seem to be much of an effect of the recipient's past responses on the overall probability of the signaller attempting another gesture.

Figure 9 illustrates the different configurations of gestures and responses that could lead a signaller to her 3rd gesture in a sequence. The probabilities of persisting after two recipient responses (or two null responses) in a row do not seem to differ from the probabilities observed in Figure 8 where the probability of persisting at each point in the sequence does not take into account the previous path of the dialogue. Figure 9 also sets forth the probabilities of the signaller's receiving a response and of using another gesture for sequences up to 3 gestures long.

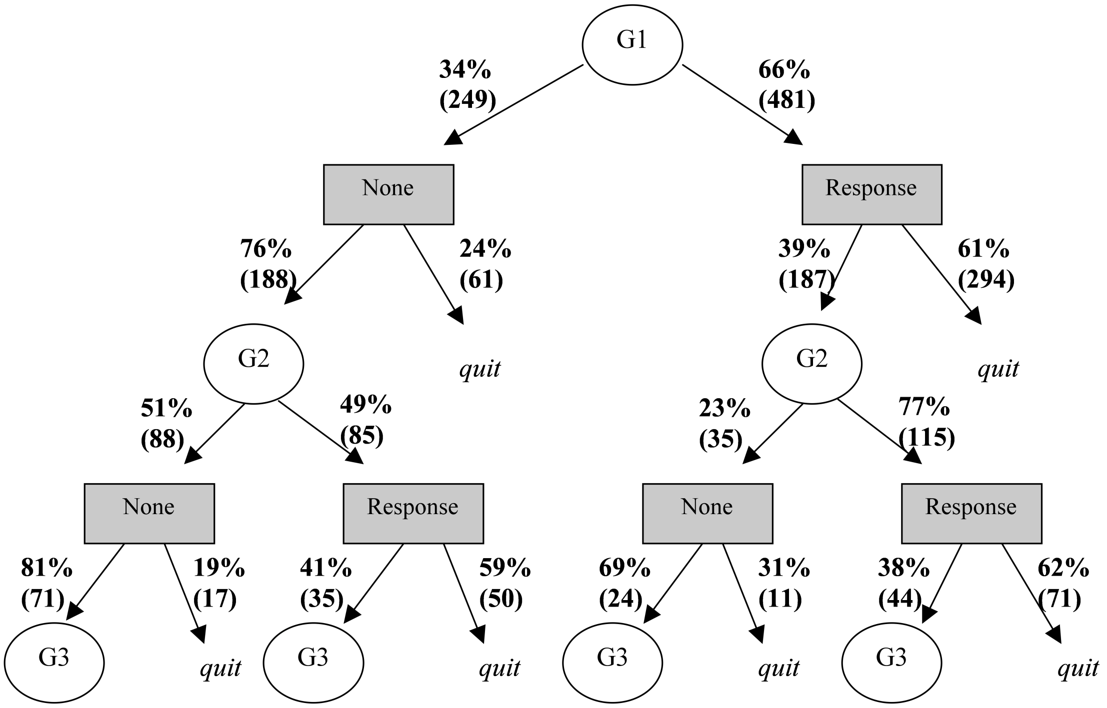


Figure 9: PROBABILITY OF CONTINUING TO GESTURE BASED ON RECIPIENT'S RESPONSES UP TO THE PRESENT.

There were 730 first-stage gestures (G1). The number of gestures decreases with every stage in the diagram because some signallers stop gesturing at each stage. The numbers in brackets indicate the number of utterances observed that matched the pattern of gestures and responses up to that point.

The diagram shows the fact that signallers are much more likely to persist following no response from the recipient regardless of whether the recipient responded to an earlier gesture or not. The low probability of a signaller's attempting another gesture if the recipient responds to the previous one reflects the likelihood that the signaller's goal was fulfilled by the response. Cases where the signaller persisted despite a response from the recipient suggest that the recipient's action was not the desired response. Responses are not all the same—many responses may fulfil the signaller's goal, whilst others may be active rejections of the signaller's request. By looking at whether or not the recipient's actions fulfil the signaller's goal, it is possible to examine how the signaller's motivation to communicate changes in response to the recipient's cooperativeness.

4.4.2 Do orangutans distinguish between a lack of response and an undesired response?

In order to investigate the signaller's likelihood of using another gesture based on whether or not their goal had been fulfilled, I analysed the probability of persisting following gestures meaning "affiliate/play." This meaning was used for analysis as it was the most common meaning of gestures and showed a high fidelity of meaning from first to second gestures in a sequence, thus raising the likelihood that "affiliate/play" gestures used as 2nd, 3rd, or 4th gestures would be parts of sequences directed towards the same meaning. The likelihood of performing another gesture was calculated for gestures in positions 1-4 in a sequence following either responses that met the goal "affiliate/play," responses that didn't meet the goal, or no response at all (Figure 10).

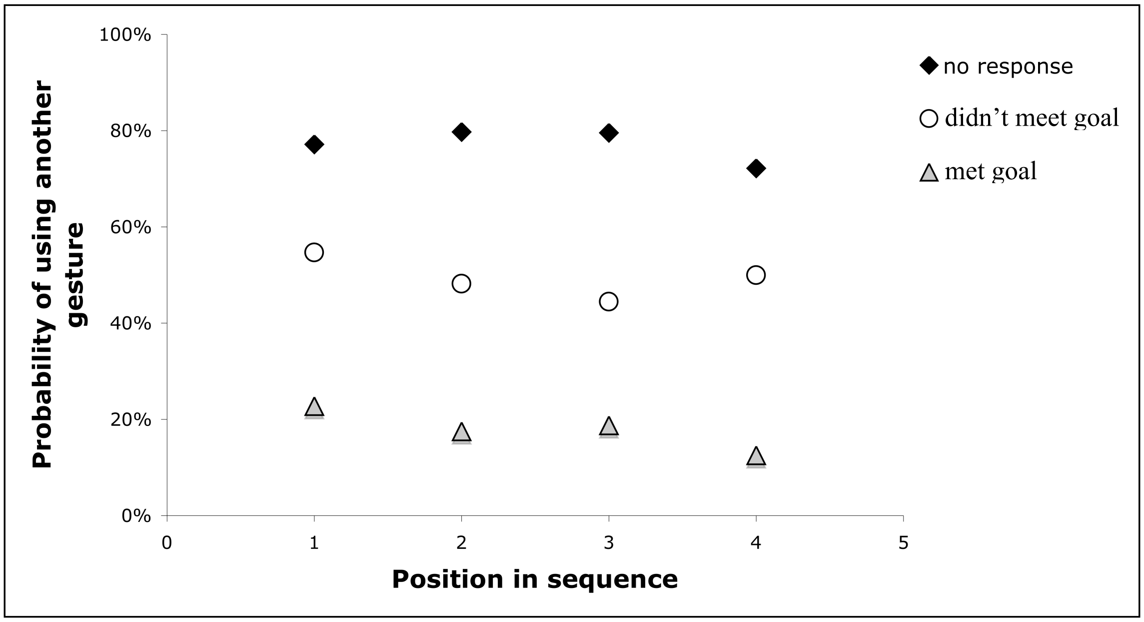


Figure 10: CHANCE OF CONTINUING TO GESTURE AT DIFFERENT POINTS IN A SEQUENCE FOLLOWING GESTURES MEANING "AFFILIATE/PLAY" BASED ON WHETHER RECIPIENT'S RESPONSE FULFILLS GOAL.

When the orangutans used gestures meaning “affiliate/play,” the chance of the signaller’s using another gesture was highest following no recipient response (77% ± 3.5%), lower following responses that did not fulfil the goal of the gesture (49% ± 4.3%), and lowest following responses that fell within the definition of affiliation or play (18% ± 4.2%). As the probability of continuing to gesture did not vary greatly over the four positions in the sequence, data from all four positions were pooled for analysis. Following no response, orangutans continued to gesture a total of 231 times and stopped gesturing 65 times. Following a response that did not meet the goal “affiliate/play,” orangutans continued to gesture a total of 136 times and stopped gesturing 125 times. Following a response that was consistent with the goal “affiliate/play,” orangutans continued to gesture a total of 29 times and stopped gesturing 114 times. These observed frequencies were compared to one another using a chi-squared test. The type of recipient response was found to significantly affect the signaller’s chance of using another gesture ($\chi^2=134.31$, df=2, $p<0.0001$).

The likelihood of a signaller’s attempting another gesture following a response that matched the meaning of the gesture was lower than the likelihood following other types of responses, but if the signaller’s goal had been fulfilled, one would expect them to never persist. The persistence rate (18% ± 4.2%) following responses that corresponded to the meaning of the previous gesture may be a product of the breadth of the category of meaning assigned to both the gestures and responses. All types of affiliation and play were combined into a single category that included actions as diverse as chasing, cuddling, wrestling, playing with objects, and sitting in contact. Perhaps the cases where the signaller continued to gesture following an affiliative response represent cases where the recipient responded generally in the right way, but did not meet the precise expectations of the signaller. If the signaller wanted to wrestle with the recipient and the recipient responded to the first gesture by touching or embracing the signaller, the signaller would likely continue to gesture in an attempt to engage the recipient in a wrestling bout.

Orangutans were more likely to continue gesturing when the recipient did not respond than when the recipient responded in a way that did not fulfil the goal of the previous gesture. Though the practical outcome of the two conditions was the same (i.e. the signaller's goal was not met), the recipient's behaviour suggests in one case that the recipient may not have perceived the gesture, and in the other that the recipient has either misunderstood the goal or rejected the signaller's attempt to affiliate. To humans, there is an important difference between a recipient who has not yet responded and one who has actively rejected a communicative attempt. Similar distinctions in the cooperative nature or willingness of the recipient have been tested using human experimenters and chimpanzee subjects. These studies have produced mixed results. Some found that chimpanzees are unable to distinguish intentions (Povinelli et al. 1998), while others suggested that chimpanzees are able to distinguish intentions, particularly in competitive contexts (Call 2003; Tomasello et al. 2003; Call et al. 2004; Hare and Tomasello 2004). The most recent studies suggest that chimpanzees have some ability to act according to the intentions of others, but there are no comparable studies of this ability in orangutans. Our finding that orangutans are less likely to persist when the recipient has responded with an action that does not meet their goal than when the recipient has not responded at all suggests that orangutans may make a distinction between an active rejection or misunderstanding and a failure to communicate.

Before any conclusions can be drawn regarding the ability of orangutans to discriminate between failed communication and communication that has not yet been successful (i.e. a lack of response), it is necessary to ensure that there is a true distinction between the two conditions. It is possible that in cases where the human observer has recorded a lack of response from the recipient, the recipient is merely responding in a more subtle way. In an observational study using conspecific recipients, it is difficult to determine whether there is actually no response from the recipient. In the present analysis, the recipient met the criteria for demonstrating no response when he continued the behaviour in which he was engaged before the gesture with no noticeable changes in attention or orientation. The term "any noticeable changes" avoided the difficulty in assessing subtle changes in temperament and gaze. It was almost impossible to determine from the video gathered whether the recipient glanced sideways at the recipient, tensed their body slightly, or otherwise acknowledged the gesture but did not respond with a large behaviour change. In order to determine whether orangutan signallers are able to determine and act in response to the recipient's understanding of their gestures, it is therefore desirable to examine sequences made by orangutans in situations where the recipient's responses can be controlled (see Chapter 5). The question of whether or not the recipient has perceived the signaller's gesture however, can be addressed by examining the signaller's use of different gestural modalities in response to the recipient's visual attention (see section 4.6).

4.5 Do orangutans persist strategically?

When a signaller's goal is not reached, she may choose to give up or to persist in trying to reach that goal. If she chooses to persist, she may employ one of many different strategies. She may repeat or exaggerate the signal that failed to receive the desired response when first used. More subtly, she may hold the final position of the gesture in expectation of a response. She may switch to another signal within the same modality, or switch to one in a different modality. Or she might choose to combine several of these strategies and hold or repeat a signal of one modality whilst adding a signal of the same or different modality (such as an *arm raise* repeated alongside a *reach*, or a *touch* held in contact with a *playface* added to it). These different strategies can be summarised as: *repeat*, *exaggerate*, *hold*, *new signal of the same modality*, *change modality*, *embellish*, and *multimodal*. This is by no means an exhaustive list of communicative strategies, but I chose to focus on these strategies because they are easily applied to gesture and focus on the signaller's choice of actions rather than the physical relationship between signaller and recipient or timing of the gestures (parameters that would lead to strategies such as *move in front of* or *speed up*). The methods of persistence chosen as strategies of interest led to an analysis that was focused on the selection and exchange of gestural units and allowed me to examine the signaller's strategies from the perspective of both morphology and discourse.

It is possible that orangutans use these strategies randomly, but it is also possible that signallers use them strategically to increase the likelihood of achieving their goal in different situations. Some types of persistence may be more effective in one condition than another. The use of different persistence strategies across different contexts may inform us as to how orangutans attempt to achieve their communicative goals under different conditions. Figure 11 illustrates the frequency of different types of persistence observed when moving from the first to the second gesture in a sequence.

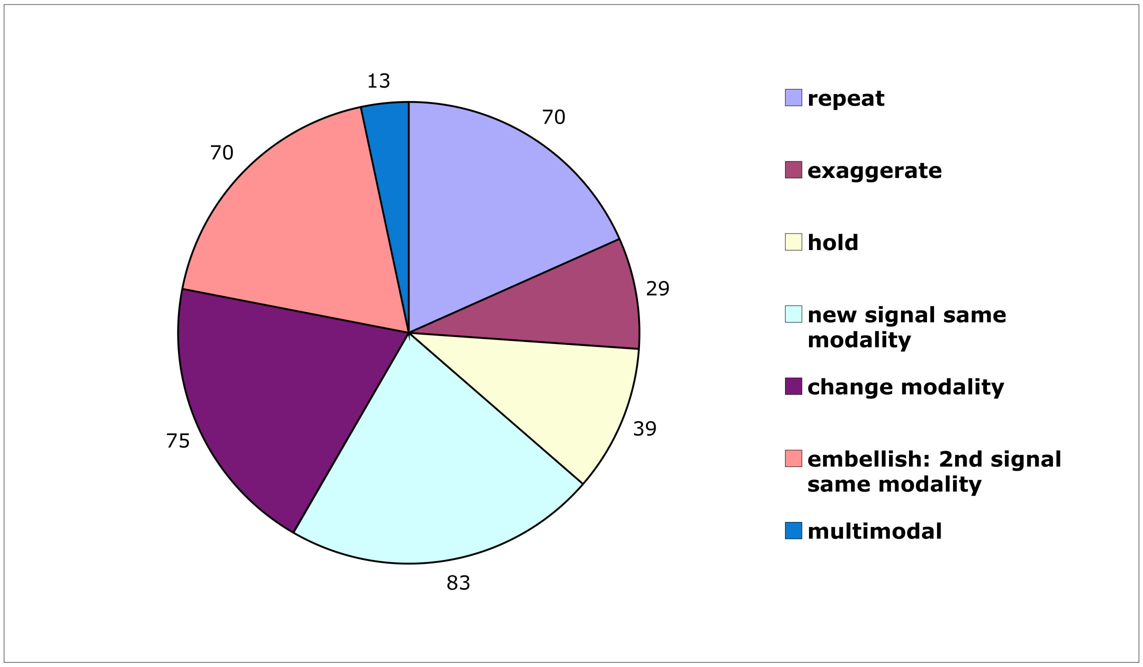


Figure 11: **FREQUENCY OF DIFFERENT TYPES OF PERSISTENCE FOLLOWING THE FIRST GESTURE.**

There were 379 cases of persistence as persisting from the first gesture by "holding" did not necessarily lead to a second gesture and thus create a sequence.

As Figure 11 illustrates, signallers persisting after the first gesture are most likely to do so by using a different signal of the same modality, and do so 22% of the time. Repeating the signal again and supplementing the original signal by holding or repeating the original signal and adding a new signal of the same modality to it both occur 18% of the time respectively. These strategies all involve using a second gesture of the same modality as the first. When they are combined with the other strategies that do not change modality (hold and exaggerate), they account for 77% of the observed 2nd gestures.

4.5.1 Use of visual versus tactile modalities

Orangutan gestures of both tactile and visual modalities can be followed by gestures of any other modality, and signallers often switch from one modality to another. However, there seems to be a hierarchy of gestural modalities. Visual gestures are more likely to be followed by tactile gestures than vice versa. If the type of persistence from the first gesture in a sequence is examined, it is clear that the probability of choosing one type of strategy over another differs depending on whether the sequence began with a visual or tactile gesture (Figure 12).

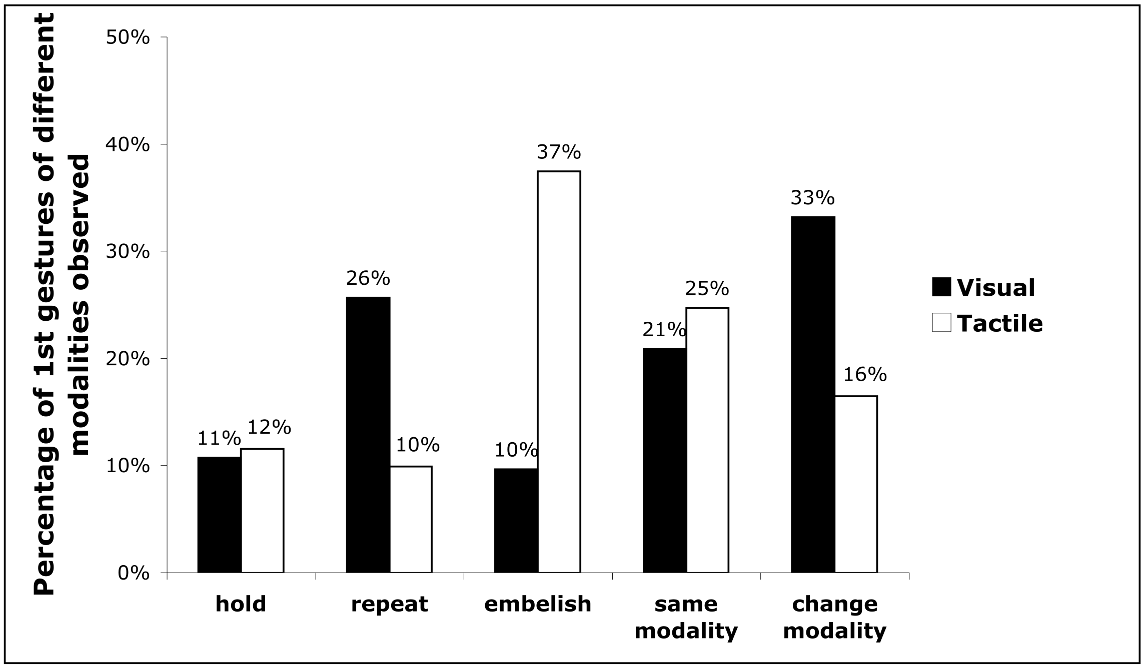


Figure 12: PERSISTENCE STRATEGIES FOLLOWING INITIAL VISUAL OR TACTILE GESTURES.

Frequency of different persistence strategies following the first gesture performed. For this analysis, single gestures were included if they were held during response waiting. Thus the total number of counts of persistence analysed was 379: 170 following visual gestures and 209 following tactile ones.

When signallers began with a visual gesture, they were twice as likely to persist by changing modality (54 out of 170 times) than they were when they began with a tactile gesture (34 out of 209 times). However, initial visual gestures were more than twice as likely (49 out of 170 times) to be repeated than tactile ones (21 out of 209 times). Sequences were most likely to be continued through embellishment (which, for this analysis, included exaggerating the initial gesture or adding a second element of the same modality to it) when they began with tactile gestures (83 out of 209 times) than when they began with visual gestures (16 out of 170 times). Chi-squared tests were performed for all of types of persistence comparing sequences beginning with tactile and visual gestures. Once P-values were Bonferroni corrected for multiple testing, significant differences were found for the persistence types *embellishment* ($\chi^2=44.60$, $p<0.05$), *change modality* ($\chi^2=12.62$, df=1, $p<0.05$), and *repeat* ($\chi^2=21.95$, df=1, $p<0.05$). No significant differences were found for the frequency of the strategies *hold* ($\chi^2=0.26$, df=1) or *new signal same modality* ($\chi^2=1.70$, df=1). In the case of tactile gestures, embellishment often took the form of the initial touch remaining in contact with the recipient while another point of contact was added (e.g. embellishing an initial *grab* by keeping hold and adding a *bite*).

4.6 Do orangutans alter their strategies according to the recipient's awareness?

Visual gestures, unlike tactile ones, might be missed by their intended recipients. Because of this, there is potential ambiguity when a recipient does not respond to a visual gesture—it may be because he is ignoring it, it may be that he did not see it, or it may be that he perceived it as having been directed towards another individual. If orangutans can distinguish between gestures that were not perceived and gestures that were ignored, one would expect that the way in which they persist following a visual gesture would depend on the likelihood that the recipient saw the gesture. If there were a chance that the recipient had not seen the signaller's initial visual gesture, then the signaller would benefit in repeating the gesture (if the recipient is now attending) or in switching to a tactile modality where she can be sure that the signal reaches its intended recipient.

Previous studies have demonstrated that great apes use the visual attention of recipients to decide how and where to communicate (Liebal et al. 2004; Poss et al. 2006) or forage (Hare et al. 2000). Results from the current study support these findings by demonstrating that orangutans were more likely to use visual gestures when they could be seen by the recipient than when the recipient was looking the other way (see Chapter 3.3). Since orangutans are able to use the recipient's present gaze in choosing modalities of gestures, it is possible that orangutans are able to take into account the gaze of the recipient in the immediate past as well. If orangutans are taking into account the possibility that their past visual gestures have not been received, one would expect that their strategies for persistence would differ based on the recipient's visual attention for sequences beginning with *visual* gestures.

4.6.1 Methodology for testing changes to persistence strategies in visual gestures

Frequencies of different persistence strategies (e.g. repeat, hold, change modality) were calculated for each of four conditions relating to the recipient's gaze:

1) The recipient is looking towards the signaller at the start of the first gesture and responds to the gesture by showing an increase in attentiveness (leaning or moving towards)
2) The recipient is looking at the start of the first gesture and continues to look in direction of recipient but does not respond by changing behaviour or showing interest
3) The recipient is not looking at the signaller at the start of the first gesture but responds by looking towards the recipient

4) The recipient is not looking at the start of the first gesture and does not respond

If orangutans are aware not only of whether or not their intended recipients can see them but also how the recipients' gaze affects whether or not they can perceive visual gestures, then their persistence strategies following visual gestures should reflect this awareness. An ability of orangutan signallers to alter their communication based on the likelihood that the recipients had witnessed a visual gesture might suggest that orangutans have some understanding of the link between seeing and knowing. Many studies have attempted to determine whether non-human primates understand that the knowledge of others can vary based on their visual experiences (Povinelli et al. 1990; Call and Tomasello 1994; Povinelli et al. 1994). Experimental manipulations have been performed to determine whether apes understand the role of the eyes in seeing or if their awareness of gaze stems from perception of body or head orientation (e.g. Povinelli et al. 1999). Most of these studies relied on heavily trained animals and unnatural experimental paradigms, an aspect of the research that was criticized as leading to over-enthusiastic interpretations of their results (Heyes 1998). In response to these criticisms, many psychologists defended the comparative approach to studying theory of mind, but acknowledged the need for more natural experimental paradigms (e.g. Gómez 1998; Matheson et al. 1998; Miles and Roberts 1998). The observational study presented here does not attempt to make any contribution to the discussion of the exact nature of orangutans' understanding of the mechanisms behind the gaze of others, but does present observational data on orangutans' use of the visual attention of conspecifics in their natural communication system. The analysis was designed to investigate whether orangutans change their communicative behaviour in a manner consistent with an understanding of whether or not their gestures could have been perceived by another individual.

4.6.2 Results: The effect of recipient's gaze on communicative persistence strategies

Do orangutans hold the final position only when the gesture has potentially been seen?

Hypothesis 1: When the recipient was not looking before the start of the first gesture, but looks at the signaller in response, the signaller should be less likely to hold the position of the first gesture than in cases when the recipient had been visually attending from the start.

When the recipient was looking during the first gesture and did not respond (but maintained gaze in the signaller's direction), orangutans held their gestures 9 times and used other strategies 79 times. When the recipient was not looking during the initial gesture but responded by looking/leaning/moving towards the signaller, signallers never held the original gesture and used other strategies on 12 occasions (Table 22).

Recipient's gaze at start of 1st gesture	**Recipient's response to 1st gesture**	**Hold final position**	**Other strategies**
Looking	No response	9	79
Not looking	Look	0	12

Table 22: FREQUENCY OF THE STRATEGY "HOLD" FOLLOWING VISUAL GESTURES AIMED AT RECIPIENTS WHO COULD OR COULD NOT SEE THE FIRST GESTURE.

Repetition is employed as a persistence strategy most often when a recipient who did not see the first gesture is now attending. However, signallers were not more likely to hold their first gestures when the recipient had been looking at the start of the gestures (χ^2=1.35, df=1, not significant). Though signallers were not statistically more likely to hold gestures when the recipient had been looking at the start, they never held gestures when the recipient could not have seen the first gesture. It is likely that the lack of significance is due to the low number of sequences beginning with visual gestures in which the recipient was not looking at the start of the first gesture but responded by looking. There were only 12 of these sequences, which is not surprising given that, in these cases, the recipient has to respond to a gesture that he has not seen by looking towards the signaller or the signaller has to move in front of the recipient before continuing to gesture. I expect that if more of these sequences were observed, the frequency of the strategy "hold" would become significant for sequences in which the recipient was looking at the start.

Are orangutans more likely to repeat or embellish gestures that the recipient did not see initially?

Hypothesis 2: If the recipient was not looking before the start of the first gesture, but looks at the signaller in response, the signaller should be more likely to repeat or embellish the original signal than she would be in cases when the recipient had been attending from the start.

If orangutans can determine whether recipients likely saw gestures that they produced, then they should be less likely to repeat signals that were perceived but failed to achieve the desired response than they are signals that may not have been fully perceived. To determine whether this was the case, I compared the frequency of strategies involving repetition of the original signal (i.e. "repeat," "exaggerate," or "embellish") made towards recipients who were looking towards the signaller during an initial visual gesture to those that were not looking (Table 23). In both cases the recipients responded to the first gesture by showing increased attention (looking if they weren't before or leaning towards if they were already looking).

When the recipient was looking during the first gesture and responded with increased attention (moving or leaning towards), orangutans used the strategies "repeat," "exaggerate," or "embellish" 34 times and used other strategies 60 times to persist in their communication. When the recipient was not looking during the initial gesture but responded by looking, leaning, or moving towards the signaller, signallers used the strategies “repeat,” “exaggerate,” or “embellish” 8 times and used other strategies 4 times.

Recipient’s gaze at start of 1st gesture	**Recipient’s response to 1st gesture**	**Elaborate (repeat, exaggerate, or embellish)**	**Other strategies**
Looking	Still looking	34	60
Not looking	Look	8	4

Table 23: FREQUENCY OF STRATEGIES INVOLVING REPETITION OF THE FIRST SIGNAL FOLLOWING VISUAL GESTURES AIMED AT RECIPIENTS WHO COULD OR COULD NOT SEE THE FIRST GESTURE.

A chi-squared test revealed a difference in the frequency of strategies involving repetition when the recipient could or could not have seen the initial gesture (χ^2=4.14, df=1, $p<0.05$). In both conditions, the recipient responded to the initial gesture with increased attention, and was looking towards the recipient at the start of the second gesture. The distinguishing variable between the two conditions appeared to be only whether or not the recipient could have seen the signaller’s initial gesture. If the recipient was likely not to have seen the first gesture, then signallers were more likely to persist by using a strategy that reiterated the first gesture (involving repetition or exaggeration).

Do orangutans change to tactile gestures more often following no response if the recipient did not see the first gesture?

Hypothesis 3: If the recipient was not looking before the start of the first gesture and does not respond, the signaller should be more likely to switch to a tactile signal than in cases where the recipient was looking but did not respond.

If orangutans can indeed distinguish between a recipient’s ignorance and rejection (i.e. not having seen a gesture versus having seen it but not responding), then they should be more likely to use a persistence strategy that ensures their second gesture is perceived in cases where the recipient may not have seen their first gesture. To determine whether this is the case, I compared the frequency of signallers switching to a tactile modality in situations where the recipient saw the first gesture but did not respond to the frequency of switching modality when the recipient did not see the first gesture and did not respond (Table 24). When the recipient was looking towards the signaller during the original visual gesture but did not respond, orangutans changed to a tactile signal 26 times and remained within the visual modality 53 times. When the recipient was not looking during the original visual gesture and did not respond, orangutans changed to tactile signals 10 times and remained in the visual modality on 16 occasions.

Recipient's gaze at start of 1st gesture	**Recipient's response to 1st gesture**	**Change modality**	**Same modality**
Looking	No response	26	53
Not looking	No response	10	16

Table 24: FREQUENCY OF CHANGING MODALITY FOLLOWING VISUAL GESTURES AIMED AT RECIPIENTS WHO COULD OR COULD NOT SEE THE FIRST GESTURE.

There was no effect of whether the recipient was initially attending or not on the likelihood of the signaller changing modality to a tactile gesture following no response from the recipient (χ^2=0.267, df=1, not significant). Signallers in both conditions changed to use a tactile signal between 33% and 38% of the time regardless of whether the recipient could have seen the initial visual gesture. This is surprising given the obvious benefit of changing to a tactile gesture if there is a possibility that the recipient did not perceive the first visual gesture and is therefore likely not to perceive the second. In these sequences where the recipient was not looking and exhibited no response however, the second gesture was also performed to a recipient who was looking away. In these cases, the signaller was performing 2 visual gestures in a row to a recipient who was not looking. Visual gestures were only used 25% of the time to recipients who were not attending (see Chapter 3.6). Thus the 16 cases where the signaller did not change modality when the recipient was not attending were extremely rare occurrences (or may have indicated that the signaller had a different motivation or could, in fact, be seen by the recipient). Tactile gestures were used more often than visual gestures whether the recipient was attending (55% of the time) or not attending (73% of the time). It is therefore not surprising that a high percentage of sequences beginning with a visual gesture included a tactile gesture as the second gesture regardless of whether the recipient was looking or not.

What is most striking is that tactile gestures as the second gesture were used very infrequently in cases where the recipient had not seen the first gesture but responded by looking towards the signaller (Figure 13).

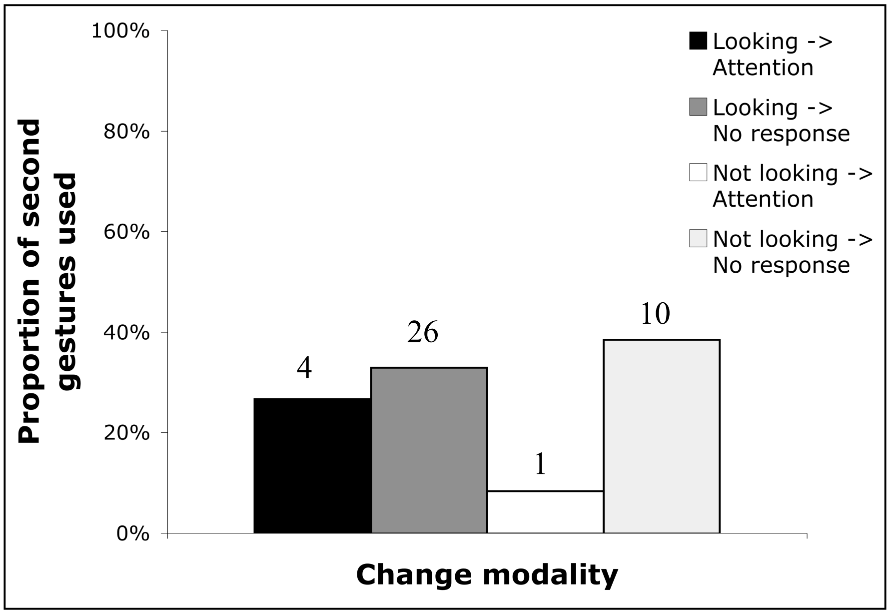


Figure 13: **FREQUENCY OF CHANGING MODALITY BASED ON RECIPIENT'S GAZE AND RESPONSE.**

The numbers above the bars are the number of times change modality was used as a persistence strategy in each of the conditions.

When the recipient had not been looking at the start of the first gesture but responded by looking towards the signaller, the signaller persisted by changing to a tactile modality only 8% of the time (a single example). The average frequency of changing modality from a visual to tactile gesture in the four different conditions (looking → increased attention, looking → no response, not looking → attention, not looking → no response) was 31% ± 13%. I tested the observed frequency of the strategy "change modality" in each of the four conditions using binomial tests and a 0.31 probability of changing modality following the first gesture. None of the observed frequencies were statistically different from the expected frequencies. The observed frequencies of changing modality were at chance levels in 3 of the four conditions, but differed from the expected value (though not significantly) in the not looking → attention condition (Binomial test, n=12, p=0.07). This finding suggests that though the recipient's gaze has no effect on the probability of changing modality in cases where the recipient does not respond, it is possible that signallers are, nevertheless, taking prior gaze into account when faced with a recipient who responds with increased attention.

4.6.3 Conclusions for visual gesture strategies: Are orangutans aware of which gestures recipients have witnessed?

The finding that signallers repeat or elaborate more often when the recipient has not seen the first gesture suggests that orangutans are capable of choosing signals by taking into account not only whether the recipient can currently see them, but also whether the recipient viewed their last gesture or not. The increase in repetitions and elaboration when recipients did not see the initial gesture also suggests that orangutans are keeping track of what gestures they have already attempted and repeat less frequently when the first gesture failed to achieve the desired response than when the first gesture was not received. The patterns of both repetition/embellishment and changing modality as strategies of persistence suggest that orangutans take into account the recipients' gaze in both the present and the past and their own past actions when choosing how to communicate.

4.7 Discussion

Orangutans use sequences of gestures to communicate about a persistent goal. Sequences beginning with gestures that were shown to have specific meanings were most likely to have second gestures of the same meaning. This finding indicated that sequences are most likely strings of gestures directed towards a persistent goal rather than pre-planned combinations of gestures that have different meanings from their component parts. Signalling orangutans were also most likely to use a new signal from the same modality when performing the second gesture in a sequence. Strategies that involved second gestures of the same modality as the first gesture made up more than 75% of all persistence strategies. These results demonstrate that both modality and meaning are conserved across the first and second gestures in orangutan's gestural sequences.

A sequence of gestures indicates that the signalling orangutan's goal has not been fulfilled. This can occur when the recipient does not respond or responds in a way that fails to meet the expectations of the signaller. From observation of the conditions under which sequences are produced, we know that orangutans persist about 75% of the time when the recipient does not respond and about 50% of the time when the recipient's response does not match the meaning of the gesture used. This finding implies that orangutans may make a distinction between their gestures not being received and being rejected or ignored.

If an initial gesture fails to achieve a desired response, there is an advantage for the signaller to persist, as the probability that a gesture would receive a positive outcome (or any outcome at all) was significantly greater than the probability that a gesture would elicit an immediate positive response. Thus, if a gesture fails to receive the desired response, it is likely that by persisting and attempting more gestures, an orangutan will eventually elicit the response she seeks.

Comparison of the different persistence strategies used when the recipient could or could not see the initial gesture demonstrated that orangutans chose signals based on the recipient's current and past visual attention. These findings suggest that orangutans monitor 1) what gestures they have already tried and 2) whether the recipient could have seen them. This implied awareness of the signaller's own past signals and the recipient's ability to perceive those signals indicate that orangutans may be sensitive not only to the recipient's receptivity, but are also tracking the recipient's reactions over time.

Though analysis of natural gesture sequences provides information about what information signallers take into account when choosing how and when to signal, there are limits to the conclusions that can be drawn using sequences directed to conspecifics. A purely observational approach is useful in determining the different types of goals that motivate the signaller and the social variables that affect the structure of sequences. However, it is difficult to control the goal of the signaller across contexts and differing social relationships, and impossible to guarantee that recipients are not responding at all even when they exhibit no gross external changes in behaviour. Since the conspecific recipients are free to respond to the signaller at any point, long sequences are very rare. Analysis of sequences directed towards conspecifics must often be limited to only the first and second gestures in order to include signals made by all individuals. Discussion of "strategies" in sequences given to conspecifics is confined to only one or two episodes of persistence in each sequence.

To investigate the communicative strategies of orangutans in more depth, it would be advantageous to have longer sequences of gestures produced under controlled conditions. Taking into account that orangutans persisted most frequently when they received no response from the recipient, a food-begging task with an unresponsive human experimenter was undertaken to experimentally elicit sequences of gestures from a subset of orangutans in the study. By presenting orangutans with an out of reach food item and a potential human helper, I anticipated that the orangutans would gesture to the experimenter and would persist in gesturing when the human did not respond. The eventual responses of the experimenter were controlled as well to simulate cases in which the orangutan successfully achieved its goal, was partially successful in achieving its goal, or failed to achieve its goal entirely. I hoped that, by controlling the recipient's response and the signaller's goal, it would be possible to identify how the orangutans' communicative strategies changed in response to the recipient's apparent understanding of their gestures.

Chapter 5: Understanding other minds

5.1 Introduction: Experimental approach to communicative strategies

Analysis of orangutan gesture sequences directed to conspecifics led to the conclusions that orangutans use sequences of gestures to persist in their attempts to reach particular goals and employ communicative strategies consonant with the interpretation that they are able to keep track of whether a recipient has seen their gestures. The data from this observational study allowed analysis of patterns of gestures produced when the recipient appeared not to meet the goals of the signaller, but was limited by an inability to control for either the signaller's goal or the recipient's response. In natural interactions, orangutans were most likely to persist in their attempts to communicate when the recipient did not respond or responded in an undesired way; in other words, when they had failed to communicate or when their requests had been refused or "misunderstood." An experimental paradigm that exploited the orangutans' tendency to persist when the recipient does not respond was employed in the hopes of eliciting sequences of gestures under more controlled conditions.

In an attempt to determine whether orangutans are, in fact, able to strategically communicate in response to the recipient's perception or understanding of their gestures, I utilized an experimental paradigm in which human recipients appeared not to understand the orangutans' requests. In this experiment, I presented captive orangutans with situations in which out-of-reach food items required human help to access but the experimenter sometimes "misunderstood" the orangutan's requests. Through the use of this paradigm, I hoped to elicit sequences of gestures that were long enough to identify potential strategic differences in the ways in which orangutans attempt to overcome misunderstanding.

5.2 Methods

5.2.1 Subjects

Subjects consisted of 3 adult female Bornean orangutans (*Pongo pygmaeus*) at Twycross Zoo, England, and 3 adult female Sumatran orangutans (*Pongo abelii*) at The Durrell Wildlife Conservation Trust, Jersey. Individuals at both zoos were housed with at least one other orangutan. All individuals had received food manually from keepers before, and most had been rewarded for returning foreign objects from their enclosures. Thus all experimental subjects had been rewarded at some point with food items and were familiar with interacting with their keepers through the enclosure bars. The head keeper for the orangutan exhibit at each zoo served as the primary experimenter, whilst I set up the experiment and recording equipment. The keeper was used as the experimenter for reasons of safety and ethics as determined by both participating zoos.

5.2.2 Design

Using a partially-modified design from Leavens et al. (2005), subjects were presented with both a highly desirable and a relatively undesirable food, allowing them the opportunity to request one or the other food by gesturing towards a human experimenter. Foods were chosen according to the husbandry practices and preferences of the orangutans at each zoo, such that all foods were familiar to the animals, the desirable food was preferred by all individuals over most other food items, and the undesirable food was typically left untouched or ignored. The head keepers chose which foods would be used based on the first foods consumed and the foods left until the end during several scheduled feedings. All food-oriented behaviour was directed towards the desirable food, implying that the “undesirable” food was, in fact, of no interest to the orangutans. Foods were, respectively, whole-grain bread and leeks at Twycross Zoo, and bananas and celery at the Durrell Wildlife Conservation Trust.

A Sony Handicam DCR TRV-38 was used to film all trials. The camera was placed on a tripod 2.0-3.5 meters from the test subject and positioned so that it filmed the actions of the subject orangutan over the experimenter's shoulder. Prior to testing, a chair was placed 80 cm from the bars of the cage. Two plastic dishes were placed on the floor, approximately 30 cm to the right and left and 30 cm forward of the chair (Figure 14). The subject orangutan was then allowed into the testing area. This involved the test subject's being separated from the rest of the group in an inside area (often a sleeping area) or called over by the keeper if the individual was tested in a larger room.

Figure 14: EXAMPLE OF EXPERIMENTAL SETUP AT TWYCROSS ZOO.

The superimposed rectangle indicates an empty food dish on the floor to the right of the chair; it was clear plastic and difficult to see.

Experiments were run at Twycross Zoo in May and October 2006 and at Durrell Wildlife Conservation Trust in July and August 2006. All trials were run prior to the afternoon or evening feed so that the individuals would be food-motivated. All individuals were separated from other adults during testing, but mothers were accompanied by their infants or sub-adult offspring; offspring present during testing ranged from 1 to 7 years of age. Feeding was delayed until the individuals had completed a trial, but no other changes to husbandry were made.

A pre-trial test using the whole-goal condition was performed once prior to experimental testing to familiarize individuals with the design and show them that it was possible to receive the desirable food. Seven females and one male were initially selected for pre-trial testing. Each was given one familiarization trial (of the whole-goal type). Two individuals did not remain attentive during this trial, moving away from the experimenter or interacting with other orangutans. These two received a second trial to determine if they would remain attentive after having seen that it was possible to receive the food. Both were included in the study after the second pre-trial phase. All eight individuals began the testing phase of the experiment, but two subjects were discarded early on after the start of testing: an adult male at the Durrell Wildlife Conservation Trust, who consistently failed to show any interest in obtaining food though he was tested in all three conditions, and one juvenile female at Twycross Zoo, who completed one experimental trial but was deemed by the keeper to be too upset at being separated from her mother.

Three experimental conditions were used. In all three conditions, a human experimenter sat facing the orangutan and remained unresponsive for 30 seconds before giving the orangutan one of the foods. The type of food delivered was the defining variable in each condition. In the non-goal condition, the orangutan was given the undesirable food. In the whole-goal condition, the orangutan was given the desirable food. In the part-goal condition, the experimenter gave the orangutan half of the desirable food and placed the remaining half out of view. Viewed as the result of communication, the three experimental conditions corresponded to the experimenter's not understanding, fully understanding, or partially understanding the goal of the orangutan.

The six adult female orangutans were tested once in each of the three conditions, giving a total of 18 trials, counterbalanced across individuals. Three experimental trials had to be rerun, one due to experimenter error and two due to the orangutan's moving out of view of the camera during the trial. Orangutans are sensitive to the presence and visual orientation of a human experimenter, communicating more often when an observer is present and favouring visual signals more when the observer can see them (Liebal et al. 2004; Liebal et al. 2006; Poss et al. 2006, also see Chapters 3 and 4). Therefore, to avoid possible effects of human visual orientation in the present study, the experimenter consistently looked towards the subject's face during the trial, but without staring into its eyes. Each subject was videotaped for 30 seconds prior to and 60 seconds after the delivery of food, and all experimenter-directed or food-directed actions were coded. Behaviour was classified as pre-delivery or post-delivery. All apparently communicative actions made by orangutans were recorded and treated as "gestures", in the broadest sense, and sub-categorized into visual, auditory, projectile, attempted barter, object retrieval, or self-directed behaviour (Table 25).

Table 25: DEFINITIONS AND CATEGORISATION OF ALL CODED ACTIONS IN EXPERIMENT.

Meanings, as determined through the study of observable goals in interactions with conspecifics (Chapters 3 and 4), are listed for gestures that appeared in both naturalistic and experimental conditions. "LF" denotes a "low fidelity" meaning (i.e. the gesture is associated with the meaning 50-69% of the time). Actions/gestures that were never produced to conspecifics are marked "N/A" in the meaning column. A "—" indicates that a gesture was used towards conspecifics but that a meaning could not be determined. Actions that appeared in interactions with conspecifics but that did not meet the criteria for intentional usage are marked "no intentional usage."

Behaviour	Definition	Meaning (with conspecifics)	Behavioural Category
Raspberry Face	Pucker lips as in raspberry sound	—	Visual
Shake object	Hold object out from body and shake	*Affiliate/Play*	
Point	One or two fingers extended during "reach"	N/A	
Reach	Hand thrust at least halfway through bars or under door	*Affiliate/Play (LF)*	
Rock/Swing	Exaggerated pendulum movement of the entire body through at least 45 degrees	No intentional usage	
Wave	Limb shaken back and forth	*Affiliate/Play*	
Wipe face	Hand is swiped across nose and mouth	No intentional usage	
Cage bang	Appendage is hit audibly against the wall, floor or climbing structure	*Affiliate/Play*	Auditory
Object bang	Object is used to perform "cage bang"	N/A	
Clap	Hands are clapped together	No intentional usage	
Kiss squeek	Sharp squeek made by sucking air through tensed lips	N/A	
Raspberry	Air is exhaled through partially tensed lips creating a buzz	No intentional usage	
Spit	Spit through bars	N/A	Projectile
Throw object	Object is thrown towards experimenter, through bars or under door	N/A	
Offer	Object is extended and held through bars or under door	*Look at object*	Attempted Barter
Fish	Object is used as a tool to reach towards one of the foods (sometimes only a few inches)	N/A	Object Retrieval
Yawn	Yawn	N/A	Self-directed

All food-directed actions, and any non-locomotor and non-functional movements (usually gestures) of the subject orangutan that took place while oriented or looking towards the experimenter, and any noise-making actions regardless of orientation, were coded from the time the experimenter sat down until he stood up (approximately 90-100 seconds). Both orangutan- and object-produced noises were counted, including blowing raspberries and banging on the cage. Necessarily repetitive actions such as rocking, waving, etc. were counted once for each sequence of movement or noise. If a pause of more than 5 seconds occurred between necessarily repetitive actions, then the continuation was counted as a new sequence.

“Repetition” was defined as consecutive reuses of a gesture or action. I counted each use of an action after the first in a bout as one repetition. Several bouts could occur in one phase of the experiment. Figure 15 shows a still frame from the video taken during one trial at Durrell Wildlife Conservation Trust during the pre-delivery phase of the experiment. The orangutan’s gesture is oriented towards the experimenter sitting in the chair to the left of the frame; the desirable food is off-screen to the lower right of the frame. Once the actions in Table 25 were coded from the video, their frequency and pattern was compared between conditions.

Figure 15: Gesture produced during testing by a subject at Durrell Wildlife Conservation Trust.

5.2.3 Procedure

The camera, chair, and empty dishes were arranged before the orangutan subject was brought into the testing area or called over to the side of the enclosure. Once the equipment was placed and the orangutan was visually attending to the experimental set-up or surrounding keeper-access area, I entered, placed a desirable food item on one dish and an undesirable food item on the other, and then left the access area. After 20 seconds, the secondary experimenter (the keeper) entered the access area and sat on the chair, facing the orangutan in a neutral position with his hands on his knees. He looked towards the orangutan, avoiding direct eye contact, and did not speak or respond to any actions of the orangutan. After 30 seconds, the experimenter delivered either the whole desirable food (whole-goal), half of the desirable food (part-goal), or the whole undesirable food (non-goal) to the orangutan. In delivering half of the desirable food, the experimenter held up the whole food item and tore it in half; then he gave one half to the orangutan and put the other half in his pocket or in his lap covered by his hands, out of view. After delivering the food, the experimenter sat down and remained in an unresponsive neutral state for 60 seconds before leaving the access area. I then re-entered the access area and removed the camera, chair, dishes, and any remaining food items.

5.3 Results

5.3.1 Inter-observer reliability

The orangutans' behaviour was analysed to determine whether they would persist in communication when their goal was not fully met, and whether they would vary their attempts following miscommunication in a way that might help a recipient determine their goal. I was the primary coder for all of the video data, but a second individual coded 16% of the total trials to examine inter-observer reliability. Agreement between the two observers was good (Cohen's Kappa = 0.85), and all discrepancies consisted of one observer failing to notice a gesture rather than disagreeing on the type of gesture.

5.3.2 Comparison of conditions before delivery of food

Since the pre-delivery phase of the experiment was the same in all conditions, the orangutans should not have been able to predict which food they would subsequently be given. Pre-delivery, neither the total number of actions nor the distribution of behavioural categories varied significantly between conditions (*Number of actions*: Friedman analysis of variance by ranks; n=6, χ_r^2=0.6, df=2, p=0.74; *Distribution of the 6 behavioural* *categories*: Friedman analysis of variance by ranks n=6, df=2 for all; for “visual” χ_r^2=4.6, p=0.10; “auditory” χ_r^2=0.9, p=0.63; “projectile” χ_r^2=0.4, p=0.82; “self-directed” χ_r^2=1.0, p=0.61; “fish” χ_r^2=3.0, p=0.22; “attempted barter” χ_r^2=2.0, p=0.37). Also, there was no difference in the number of gesture types used pre-delivery for any of the three conditions (Friedman analysis of variance by ranks n=6, df=2, χ_r^2=4.1, p=0.13). Since pre-delivery behaviour did not differ between conditions, we conclude that conditions did not differ in any way other than the type of food delivered to the orangutan.

5.3.3 Do orangutans communicate strategically in response to others' apparent understanding of their gestures?

If orangutans have a specific goal in mind when attempting to communicate, then they should cease signalling if their goal is reached (McFarland and Bösser 1993; Shettleworth 2001). I compared the total number of gestures an orangutan made post-delivery in each of the three conditions to determine whether they stopped signalling once they had received the desired food item. The number of gestures varied significantly with experimental condition (Friedman analysis of variance by ranks; n=6, χ_r^2=11.6, df=2, p=0.003). In the whole-goal condition, when the entire desirable food was delivered, all but one orangutan ceased signalling entirely and several individuals retreated into their cages, breaking off contact with the experimenter; only one individual persisted in signalling, making a single barter attempt. In both the part-goal and non-goal conditions, all of the subjects continued to signal to the experimenter after the delivery of food. The orangutans used significantly more gestures following delivery of half of the desirable food than after delivery of the undesirable food (Wilcoxon signed ranks test, z=-2.04, p=0.04). Since the orangutans signalled more frequently in the part-goal than in the non-goal condition, the cessation of communicative behaviour following delivery of the whole desirable food is unlikely to be an artefact of food processing (e.g. consumption temporarily suppressing other activity).

Use of repetition of gestures

In gesture sequences made to conspecifics, the second gestures made by orangutan signallers are repetitions of the first gesture between 18% (Chapter 4) and 50% (Liebal et al. 2006) of the time. In the observational study of sequences, I found that orangutans were more likely to repeat gestures if the recipient had not seen their first attempt than if the recipient had seen the first gesture. This finding suggested that orangutans might use repetition when they are unsure whether their gestures have been received. In the experimental paradigm, the three conditions simulate different levels of communicative success. If orangutans are using repetition communicatively and keeping track of the success rate of their signals, one might expect the frequency of repetition to vary with the degree to which their goal has been met.

I compared the frequency of repetitions (i.e. repeating a gesture immediately after it has been used already) for each individual between conditions (Figure 16). After delivery of food, the numbers of repeated gestures differed between experimental conditions (Friedman analysis of variance by ranks; $n=6$, $\chi_r^2=11.2$, $df=2$, $p=0.004$). When an undesirable food was given, orangutans were significantly less likely to repeat gestures than they were after delivery of half of the desirable food item (Wilcoxon signed ranks test, $z=-2.21$, $p=0.03$). The difference in tendency to repeat gestures shows that orangutans are able to evaluate their own level of success in communication, and modify their subsequent attempts accordingly. If they perceived their prior efforts as having partially succeeded or simply not-yet-failed to communicate their meaning, the orangutans in our study repeated signals up to half the time. In contrast, if they completely failed, they avoided repetition.

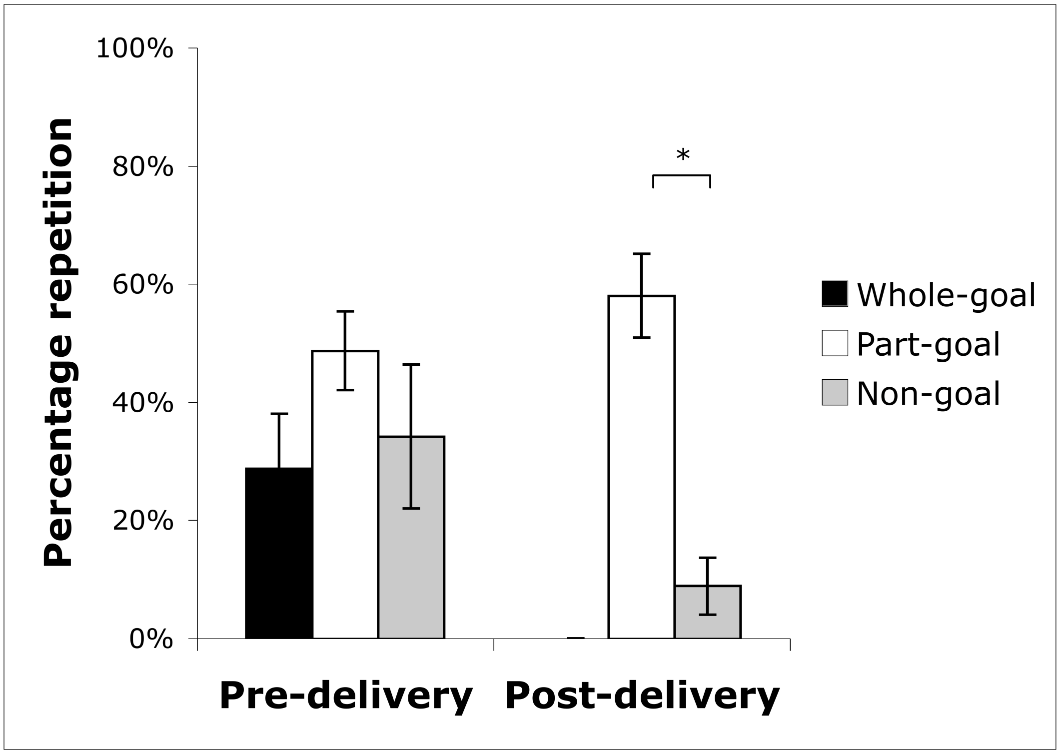


Figure 16: REPEATED COMMUNICATIVE BEHAVIOUR.

Mean percentage of all behaviour that comprised repetition, in the pre-delivery phase and post-delivery phases of the experiment.

It could be argued that repetition might stem simply from increasing frustration over time (Roberts 1981; Dickinson and Balleine 1994; Stout et al. 2003). However, the latency between communicative attempts tended to increase over time as recipients failed to respond; showing that, as they continued to receive no response, orangutans lost interest rather than becoming increasingly frustrated (Spearman's rank correlation, r=0.30, p=0.002; data from pre-delivery phase used in order to include all three conditions). Gestural repetition therefore cannot simply be attributed to frustration, but reflects an accurate understanding by orangutans of the relationship between their attempts to communicate and their level of success.

Use of different gestures

If orangutans realize that their signals have been entirely ineffective at achieving their communicative goal, then they should switch to other signals rather than persisting with those that have failed (Bates et al. 1979). Consistent with this, I found that in the non-goal condition, when an undesirable food was delivered, orangutans used more types of gesture than in the part-goal condition, when they received part of the desirable food (Wilcoxon signed ranks test n=6, z=-2.04, p=0.04). I also calculated the frequency of gestures that were used for the first time in the post-delivery phase ("novel" gestures) by each individual, in both the part-goal and non-goal conditions. Novel gesture types were more frequent in the non-goal condition (Wilcoxon signed ranks test n=6, z= -2.06, p=0.04; see Figure 17); in the part-goal condition, gestures were likely to recur, i.e. the same gesture was used pre- and post-delivery. Note that, although these data show that gestures found ineffective by orangutans during our experiment tended to be discarded, the same gestures fail to receive a response the majority of the time they are used in daily interaction with human keepers, yet remain in the orangutans' active repertoire. The changes in gesture use are temporary, forming part and parcel of each communicative exchange.

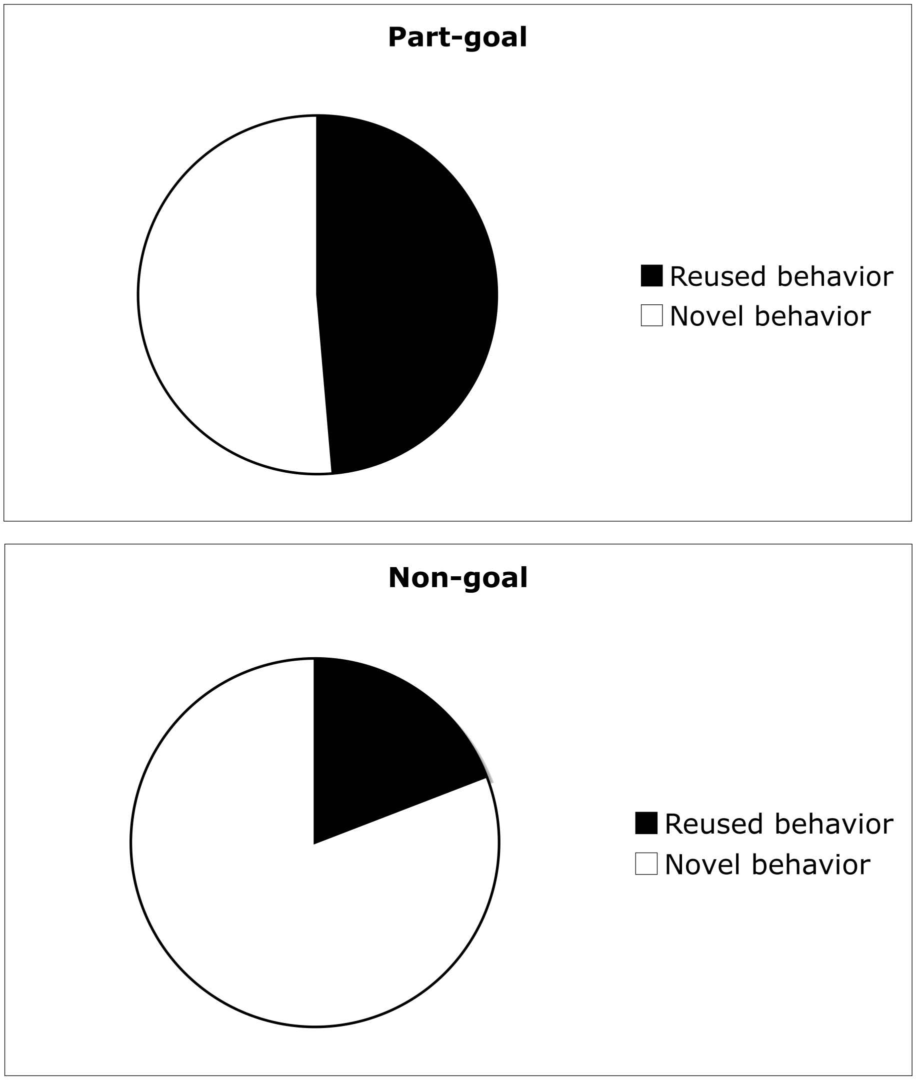


Figure 17: USE OF NOVEL BEHAVIOUR.

Average percentage of "novel" behaviour (i.e. actions used post-delivery but not pre-delivery) used in the post-delivery phase of the experiment, in non-goal and part-goal conditions.

5.4 Meaning of gestures

Only 10 of the 17 actions elicited in the experimental design also occurred in interactions with conspecifics. Of these 10, only 6 were used as intentional gestures with conspecifics, and only 5 of those could be determined to reliably predict signallers' observable goals. The gestures that appeared in both experimental and natural communicative attempts were *raspberry face (hit ground/object)*, and *offer*. Four of these gestures (*shake object*, *reach*, *wave*, *cage bang*) were used to initiate affiliation or play in conspecific interactions. The gesture *offer* seemed to be used with conspecifics to draw another's attention to an object that the signaller was holding. *Raspberry face* met the criteria for use as an intentional gesture, but there were not enough examples of it to determine whether it had a particular function. The four "affiliate/play" gestures seen in both contexts were particularly salient examples of the set of gestures associated with the affiliate/play meaning in conspecific interactions. It is possible that these gestures are used when the signaller is particularly excited. This could help to explain their occurrence in both play and food-begging contexts. Lending support to the excitement hypothesis; *shake object* was observed during display behaviour by both males and females, but it was typically directed towards visitors at the zoos or did not have a clear recipient and so was not included in the study of conspecific gesture. It is also possible that the artificially-imposed distance between signaller and recipient in the experimental design meant that the signaller required the cooperation of the experimenter to fulfil their goal of retrieving the food, thus causing the signaller to use affiliative signals in her attempt to elicit aid. Though the end result in both the cooperative experimental and food-sharing natural scenarios is the same, the signaller may make a distinction between requesting aid from a keeper and food-sharing from a conspecific and subsequently choose different types of gestures.

The different types of gestures used by the orangutans in experimental and natural settings unfortunately meant that direct comparison of strategy in the two conditions was not possible. The strategies orangutans used when communicating with conspecifics seemed to be aimed at reiterating or emphasising their goals when recipients did not respond as desired, and appeared to take into account whether recipients could have seen the past gestures or not. The strategies orangutans used in the experimental setting appear instead to focus attracting the attention of the experimenter (who remained unresponsive for long periods of time) and differed based on how close the experimenter came to fulfilling the signaller's goal.

5.5 Discussion

The sequences of gestures produced by the same six orangutans under natural circumstances to conspecifics demonstrated both persistence and elaboration (see Chapter 4). However, it was difficult to determine what the specific goal of the signaller was in each situation and whether the recipient's subsequent actions represented misunderstanding, refusal, or response to each gesture. The experimental design allowed us to control for both the goal of the signaller and the response of the recipient. In addition, the periods during which the experimenter was unresponsive elicited extremely long sequences of gestures from the orangutans, which allowed observation and analysis of more complex patterns of gestures. Rather than relying on a single transition from the first to second gesture to gain insight into persistence strategies, the sequences produced by the six experimental subjects often provided 5 or more transitions to analyse. The sequences produced during the three experimental pre-delivery phases and the two post-delivery phases that involved persistence ranged from 1 to 25 gestures long, but on average contained slightly over 6 gestures (mean 6.39 ± 5.01).

The six orangutans tested appeared to have a specific goal in mind, i.e. gaining a desirable food item, which they attempted to achieve indirectly by communicating with the human experimenter. As they did in conspecific interactions (Chapter 3), the orangutans in the experiment persisted in their communicative attempts when their goal was not met. This effect of persistence was also found in the chimpanzees tested in the Leavens et al. (2005). Orangutans have previously been found to be sensitive to the visual attention of an interlocutor in experimental settings (Liebal et al. 2004; Poss et al. 2006), but the orangutans in the current experiment went much further, acting on the apparent understanding of the recipient. The orangutans in the present study appeared to distinguish between being partially understood (when given part of the desired food) and being completely misunderstood (when given an unwanted food item). Their subsequent communicative attempts reflected this distinction. When confronted with a response that suggested partial understanding of their desire, the orangutans continued to use those signals they had used before the delivery of any food, often giving a signal repeatedly. When they were given the wrong item altogether, they instead chose to use other signals rather than to repeat those used already, and they avoided signal repetition, often attempting each new behaviour only once. The combined results of the communicative strategies identified in natural and experimental conditions suggest that orangutans are highly sensitive to the behaviour of their intended recipients and alter their communicative attempts in a manner consistent with an understanding of others as autonomous, intentional beings.

The strategy employed by the orangutans under experimental conditions resembles that of the parlour game “charades,” in which the player tries to get her team to guess the name of a book or movie by acting it out non-verbally. As the player gesticulates on stage, her team calls out their guesses as to what is being portrayed. If your team is close to your answer, the best strategy is to repeat and refine the signals that have already partially worked. But if your team completely misunderstands your gestures, it is better to switch to new signals until they guess something close to your goal. This strategy not only maximizes a player’s efficiency in choosing an effective indicator, but also communicates to the team how close they are to understanding the intended meaning. While the communication sequences of the orangutans here are perhaps not as sophisticated, they nonetheless accomplish the same objectives. By maximizing efficiency at searching for an understood signal and homing in on those that achieve partial success, orangutans are able to overcome misunderstandings. In the absence of a shared lexicon, one way of arriving at a shared meaning is to adopt a charades-like strategy, transmitting not only the content of the intended message but also a signal indicating how well you have been understood. If the recipient can use this information, then the signaller and recipient will be able to arrive at a common understanding much faster. This strategy offers one possible pathway toward constructing a shared lexicon from learned or ritualized signals. Though there is no evidence that meaningful orangutan gestures are acquired through this process, investigation into the strategies employed in the intentional communication of apes may nevertheless provide insight into the pre-linguistic devices that helped construct the very earliest forms of hominin language.

Chapter 6: Maternal influence on infant behaviour and gesture

6.1 Introduction

Though orangutans use gestures in flexible and strategic ways, suggesting they view other individuals as autonomous agents capable of having both intentions and perceptions, there is little indication that the forms of their gestures are very flexible. Very few group-specific gestures have been identified in any ape species (Pika et al. 2003; Pika et al. 2005; Liebal et al. 2006); and if gestures were socially-learned, one would expect a high variation in the forms or use of gestures between groups. Faced with a lack of evidence that gestures are culturally-transmitted, Call and Tomasello posit that gestures are acquired through the "ontogenetic ritualisation" of movements used in social interactions (2007). In this process, a movement such as shoving another out of the way becomes ritualised into a *nudge* as an individual learns (through association) that only the beginning of the movement is necessary to achieve the desired result. During this association of the first "intention movement" of an action with the recipient's response, the actor/signaller does not need to understand the process or make any cognitive leap to using a signal to indicate intention to another rather than an action to manipulate another. The process of ontogenetic ritualisation is a likely candidate for many of the gestures that resemble the functional movements that would bring about the desired results, but it is difficult to attribute all of the gestures I observed in this study to this process. Given how sensitive orangutans' use of gesture is to the actions and attention of others, it seems unlikely that the forms of individual gestures are not socially influenced by the gestures of others during development. In an attempt to determine whether social learning could play a role in the acquisition of gestures, I investigated the social influence of others on the behaviour and gestures of orangutan infants, focusing on the social partner that is the most important to the infant during development, the mother. I was also interested in determining the extent to which orangutan mothers shape the actions and knowledge of their infants, particularly when they are housed socially with other orangutans. When infants are given the choice of many social partners, how much influence over the infant does the mother retain?

Primates have longer juvenile periods and inter-birth intervals than comparable non-primate mammals (Harvey and Clutton-Brock 1985), and so mother-infant relationships have a correspondingly greater potential for influencing primate behavioural development. Amongst primates, juvenile phases and inter-birth intervals are longest in great apes, and greatest of all in orangutans (Galdikas and Wood 1990; Wich et al. 2004). Orangutans are characteristically solitary, and juveniles typically disperse when their mother has another infant (van Noordwijk and van Schaik 2005). An infant orangutan remains with its mother for up to 8-10 years, and for most of this time the infant is primarily responsible for maintaining proximity to its mother (van Adrichem et al. 2006). During all this time the mother provides the primary, if not the only, model of social and ecological competence (Galdikas and Wood 1990; Wich et al. 2004). Thus the potential exists for substantial maternal influence on an infant orangutan's development, and for the vertical transfer of information critical to survival.

Several lines of evidence have led researchers to suggest that observation of the mother is important in great ape development, because it allows social learning. In chimpanzees, though there is no evidence of active teaching of tool use, mothers are tolerant of infants observing them even when the infant's observation hinders the mother's actions (Hirata and Celli 2003; Lonsdorf 2006). Infant chimpanzees choose preferentially to observe individuals of the same age or older, suggesting that they select models likely to be more competent than themselves (Biro et al. 2003). In a study of nut-cracking, Biro et al. (2003) found a high correspondence in laterality between chimpanzee siblings, which they attributed to copying from the mother; but conversely, Byrne and Byrne (1991) found no tendency in mountain gorillas for hand preferences to run in families in several skilled food-preparation tasks. Lonsdorf (2006) found that chimpanzee infants whose mothers spent more time termite-fishing acquired elements of this skill earlier. Again, the effect was attributed to social learning, but differential opportunities for practice may also have mediated the variation. The geographical distribution of some chimpanzee habits is consistent with spread by social transmission, and not easily understood otherwise (McGrew et al. 1997). In orangutans, the evidence for ‘cultural’ transmission of habits is even stronger (van Schaik et al. 2003; Byrne, 2007). Orangutan infants in the wild spend longest watching their mother process foods that require the most steps to process (Jaeggi 2006), which suggests motivation to acquire skill by watching; although these same foods may be the least available to young orangutans and thus particularly desired. Studies of rehabilitant populations have found that orangutans are able to acquire novel behaviour by imitation, and actively prefer to observe models with whom they have affiliative relationships, such as caretakers or older siblings (Russon 1996; Stoinski and Whiten 2003; van Schaik et al. 2003; Russon et al. 2007).

Surprisingly little is known, however, of the mechanisms by which young apes are influenced by others. Does the mother form the primary model for an infant great ape only when no other is available, or do infants exert active choice? Behavioural synchrony between mother and infant is sometimes observed; but is this merely a result of shared environmental factors, or is synchrony actively managed—and if so, by whom? Does attention to the mother or another skilful adult routinely mediate the course of development of an infant great ape, or are the existing reports of social learning isolated special cases? Working with a captive population presents advantages in addressing these questions, as infants have unrestricted access to other adults and juveniles as potential models, and observation can be detailed and prolonged.

I conducted three observational studies aimed at investigating the immediate and lasting influence of the mother on the infant. Study 1 investigated the occurrence of behavioural synchrony between mother and infant orangutans, focusing on directed behaviour similar in both form and goal. Study 2 aimed to determine whether bouts of synchrony were a result of coincidence or actively managed. Study 3 compared gestural repertoires of infants with those of their mothers and other adult orangutans, with the aim of assessing whether the mother's behaviour has a lasting impact on the repertoire of the infant.

6.2 Data Collection

6.2.1 Subjects

Seven infant orangutans, aged between 10 and 25 months (mean 17.6 ± 5.5) were included in the study. The sample comprised five female and two male infants hosed across the three zoos (Apenheul, Durrell, and Twycross). The mothers ranged in age from 11 to 42 years. Six infants were mother-reared, but one mother had, in addition, adopted a second, unrelated infant (a female), which was included in the study. One juvenile (a female, aged between 30 and 48 months over the course of the study) was added to the sample in study 3 because she had also been adopted by an unrelated female. The youngest female in the sample was primiparous, and the female who adopted the infant had suffered a still birth but had subsequently adopted the above mentioned juvenile and so had experienced both pregnancy and infant rearing, though not on the same infant. Four infants (including the adopted infant) and the adopted juvenile were Bornean orangutans (*Pongo pygmaeus*) and three were Sumatran orangutans (*Pongo abelii*). Husbandry remained unaltered throughout the studies and no training was required. All observations were made from public access areas, though often during times when the exhibits were closed to the public.

6.3 Study 1: Behavioural Synchrony

In order to identify opportunities for possible social learning, I investigated the degree of behavioural synchrony between mother and infant during actions that were directed towards objects, individuals, or locations. The target as well as the action was taken into account such that object manipulation could be counted as behaviourally synchronous only if it involved the same object or objects of similar types (e.g. both playing with sawdust). Synchronous, directed action involves one individual matching the location and type of another's behaviour. This matching need not be the product of true imitation of another's actions, but might involve other learning mechanisms such as emulation or stimulus enhancement (see Whiten and Ham 1992; Zentall 1996 for reviews). Regardless of the particular mechanisms involved, synchronous actions signal moments when one individual is likely attuned to the behaviour of another and changes their own behaviour in response to some aspect of the other's actions. Synchrony might also arise through coincidental simultaneous execution of similar actions, particularly if governed by the presence of a common stimulus (such as the presence of food). However, if coincidental synchrony is minimised by avoiding periods of time when individuals are both exposed to particularly salient external stimuli, the remaining cases of synchrony are ideal candidates for the investigation of social influence and learning.

6.3.1 Procedure

Data were collected during 3 visits to the zoos over the course of 20 months: Twycross April-June and October 2006, Durrell June-September 2006, and Apenheul November 2006. I recorded the behaviour, gaze, and proximity of each mother-infant dyad using scan sampling at one-minute intervals, avoiding scheduled periods of feeding so as to minimise environmentally-determined similarity of behaviour. I defined the gaze of an individual as whether it could easily view the other without having to turn its head (see Chapter 2). In practice, this meant that each individual was defined as having about an 80-degree field of vision. When the direction of gaze was able to be determined on the video from orangutans' eyes, the direction of gaze was the centre of the field of vision. When the direction of the eyes was not clearly visible, the front of the face was used as the centre of the field of vision. If it was possible to determine that an individual's eyes were shut or their gaze was blocked by an object, then their recorded field of vision was correspondingly reduced. Mother and infant were considered to be behaviourally synchronous if they performed equivalent actions during the same scan involving a similar object or location. Only those actions done independently by each individual (i.e., excluding joint physical play, holding hands, and co-locomotion) were considered as potentially synchronous. These criteria were intended to limit analysis to those actions that both parties could begin and end autonomously and that were unlikely to be independent reactions to external events whose concurrence was coincidental.

6.3.2 Results

A total of 3309 scans were made (between 450 and 600 per dyad), of which 146 were made alongside a second observer in order to establish inter-observer reliability. Reliability was assessed for each variable measured using Cohen's kappa. Kappa values ranged from 0.56 for "outcome," to 0.94 for "direction" (mean 0.78 ± 0.14). These values indicated that the strength of agreement for all variables was either "substantial" or "almost perfect" (Landis and Koch 1977).

The mean frequency of behavioural synchrony for all dyads was 4.5% of the time budget (± 3.0%, range from 1.3 to 9.9 %). When non-directed actions and actions that did not have the potential to be performed independently by both mother and infant (i.e. locomoting, resting, sleeping, nursing) were excluded, the frequency of synchrony increased. When the number of synchronous actions were compared to the number of potentially-synchronous actions performed by the mother, the frequency increased to 8.7% (± 4.4%). When synchronous actions were compared to the infant's potentially-synchronous actions, the frequency was 6.4% (± 4.2%). The rate of synchrony derived from the frequency of potentially-synchronous actions in both mother and infant was 7.6% (± 4.1%). Of the 177 observed cases of synchrony, 77% were actions directed towards objects or individuals outside the mother-infant dyad. The remaining 23% were synchronous actions that occurred during affiliation or play, such as biting each other's hands or executing a headstand at the same time.

Analysis of proximity data from all dyads revealed a significant difference between mean proximity of mother and infant dyads during synchronous and non-synchronous behaviour (Paired t-test: $t_6 = 3.44$, $P = 0.014$, two-tailed probability). Mean proximities for the seven mother-infant dyads during non-synchronous actions ranged from 0.44 to 3.95 m, (mean = 2.28 ± 1.25), and from 0.40 to 1.17 m during behavioural synchrony (mean = 0.77 ± 0.51).

Using the gaze direction of both mother and infant, each scan was scored according to whether each individual could potentially see the other. Mothers had infants within their line of sight for a greater percentage of synchronous than baseline behaviour ($\text{mean}_{\text{synchrony}} = 84.5 \pm 8.0\%$, $\text{mean}_{\text{baseline}} = 44.2 \pm 5.5\%$). Similarly, infants could see their mothers more often during synchronous than baseline behaviour ($\text{mean}_{\text{synchrony}} = 84.5 \pm 9.2\%$, $\text{mean}_{\text{baseline}} = 52.6 \pm 7.7\%$). Paired t-tests revealed that both mother and infant could see the other significantly more often during synchronous actions (mothers: $t_5 = -9.31$, $P = 0.0007$, two-tailed probability; infants $t_5 = -5.87$, $P = 0.004$, two-tailed probability; Figure 18).

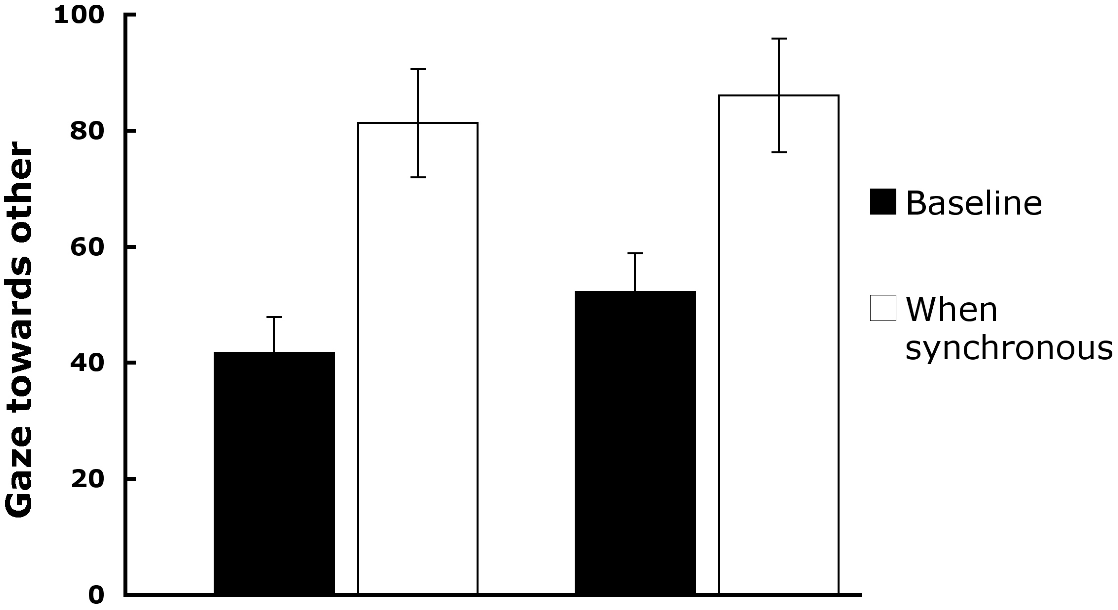


Figure 18: Mean percentage of one animal being able to see the other during baseline (N =3743) and synchronous (N = 176) behaviour of mother and infant.

Data from the 7 dyads were pooled to obtain mean values. Standard deviation is displayed as error bars.

6.3.3 Discussion

In the seven mother-infant dyads, around 5% of all actions were synchronous directed behaviour; identical in both the type of action performed and the target of the action. During periods of synchrony, the proximity of mother and infant and the potential for each to see the other were increased in all dyads. Synchronicity varied widely between dyads. This variability may have been a product of husbandry, as dyads at the same zoo had similar rates of synchrony. Variables such as enclosure size, access to objects or foraging substrates, or the number of other individuals with which the pair can socialize may have affected opportunities or motivation for behavioural synchrony. Synchrony was unusually low for the two infants cared for by the same female (1.3% for each). This may have been because the two infants were influenced by each other's behaviour, or because they shared their enclosure with the greatest number of other individuals.

Correlation between proximity and synchrony might reflect one individual's matching the other's behaviour—for example, an infant might tend to copy its mother—but the correlation could also be an artefact of exposure to more similar external stimuli when animals are close together. In that case, each party should be equally likely to cause synchrony by switching to the other's behaviour. Study 2 was designed to determine whether one individual in particular is responsible for matching the behaviour of the other, and if so, which.

6.4 Study 2: Responsibility for synchrony

Continuous focal sampling was used to determine whether either the mother or infant had a predominant role in creating synchrony. It was possible that the higher activity budgets typical of infants might cause them to switch behaviour more frequently, thus spuriously appearing to have synchronised with their mothers more often than their mothers synchronise with them. I therefore examined, in addition, those changes of behaviour that ended synchronous behaviour by switching to a new action.

6.4.1 Procedure

Three Sumatran mother-infant dyads (*Pongo abelii*) housed at Durrell Wildlife Conservation Trust were observed for this study, during June-September 2006. An Olympus digital voice recorder (VN-1000) was used to record the behavioural changes of both mother and infant during 10-minute sessions of continuous focal sampling. Number of observation sessions per dyad ranged from 13-22 (mean 18.33 ± 4.73).

6.4.2 Results

The three mother-infant dyads gave a total of 551 minutes observation, with time per dyad ranging from 126 to 219 minutes. An index of responsibility (Hinde and Spencer-Booth 1967) was calculated for each mother-infant pair, in order to find out which individual was predominantly responsible for synchrony. This index was calculated according to the following formula:

$$\frac{\text{\# synchronous events created by infant}}{\text{all synchronous events created}} - \frac{\text{\# synchronous events ended by infant}}{\text{all synchronous events ended}}$$

The index of responsibility was positive for all three dyads (0.39, 0.29, and 0.10), indicating that in each pair, the infant had a greater role in creating and maintaining synchrony. In addition, I computed the relative frequencies of making and breaking synchrony by the infant rather than the mother, for each dyad. In two of the three dyads, the infant was significantly more likely than the mother to create synchrony, and in the third case the trend was in the same direction (Sign test: *dyad 1*: $P = 0.043$, $N = 25$; *dyad 2*: $P = 0.016$, $N = 7$; *dyad 3*: $P = 0.08$, $N=21$; all tests two-tailed). In none of the three dyads was one party significantly more responsible for ending synchronous events than the other (Sign test: dyad 1: $P = 0.38$, $N = 25$; dyad 2: $P = 0.15$, $N = 7$; dyad 3: $P = 0.45$, $N = 21$; all tests two tailed).

6.4.3 Discussion

Infants were responsible for matching their mother's behaviour significantly more often than the converse, but were no more likely to end synchronicity than were their mothers. As a consequence, infants had the greater role in creating and maintaining behavioural synchrony. This may reflect a tendency in infant orangutans to copy the mother's actions and thus learn socially from her. That hypothesis is supported by the finding, in the first study, that both parties were particularly close and particularly able to watch each other during synchrony. These results do not show, however, that this maternal influence has any lasting effect on the infant. Study 3 was therefore designed to explore whether maternal influence had any developmental consequences, by examining the overlap of mothers' and infants' gestural repertoires.

6.5 Study 3: Maternal influence on gesture

If social learning plays any role in the ontogeny of an orangutan's repertoire of communicative gestures, the mother would be the most likely social influence. In this case, one would expect that the gestures of young orangutans would more closely resemble those of their mothers than those of other orangutans. Conversely, if all gestures of young orangutans are acquired individually (for instance, by ontogenetic ritualisation from non-communicative actions, through unintentional shaping during interactions: (see Call and Tomasello 2007), an infant's gestures would be no more likely to match those of its mother than those of any other orangutan. To test whether the mother does influence an infant's gestural repertoire, I compared the overlap of each infant's gestural repertoire to that of its mother and to those of other adult females in their group. Only adult females were used for comparison, because ape gestural repertoire is known to vary widely with age, so the gestures of animals within the same age class are likely to be similar (Tomasello et al. 1989). If association with the mother influences the infant's development or expression of gestures, its repertoire should more closely resemble its mother's than those of other adults. To ensure that any heightened similarity between infant and mother was an effect of time spent in association rather than any heritable predisposition towards developing or using certain types of gestures, I compared the gestures of an adopted infant and juvenile to those of their adoptive versus their natural mother. This comparison also allowed me to examine whether infants reared by unrelated females were any less influenced by their mothers' behaviour than were infants raised by their biological mothers.

6.5.1 Procedure

Using a Sony Handicam DCR TRV-38, video was collected at the three zoos included in study 1, during the same observation periods, focusing on social interactions likely to involve gesture (i.e. active social interactions such as locomotion, play, object manipulation, or foraging). Video collected during an earlier visit to Apenheul from March to May 2005 was included in this study in order to generate more comprehensive gestural repertoires for the adult females and juvenile observed in this study.

Movements of the head, limbs, or body that were “physically ineffective” (i.e., did not accomplish a physical goal such as locomotion or scratching) and occurred in the presence of another individual were treated as gestures. Cases of gesture use were extracted as clips and coded for analysis. The repertoires of seven infants (5 females and 2 males) were compared to their mothers’ repertoires and to the repertoires of other adult females (aged 13-45) in their group who had also had offspring. One adult female at Apenheul was not included in the analysis as she had never had an infant and was primarily housed by herself. Mothers were considered to be the adult female who cared for the infant, regardless of genetic relation, as one of the infants had been adopted a few days after birth by another female in her group. Gestures common to an infant’s repertoire and that of a female in its group could potentially have been learned by watching the other individual. To estimate the potential contribution of learning from the mother, I compared the overlap in repertoire between infant and mother with that between the same infant and other females.

Additionally, I compared the repertoires of two adopted orangutans, an infant and a juvenile, both female *Pongo pygmaeus*, to those of their biological and adoptive mothers. Both immatures had the same birth mother and adoptive mother and all four animals shared an enclosure at Apenheul Primate Park. This comparison was intended to find out whether the gestural repertoire of an adopted orangutan is more influenced by inheritance or upbringing, and whether adoptive mothers have a comparable influence to natural mothers.

6.5.2 Results

From a total of 1581 potential gestures coded, I compiled repertoires for all individuals. I determined the number of gestures that each infant shared with each adult female (who had reproduced) within their groups. I calculated a measure of repertoire overlap that took into account the number of shared gestures (the overlap) and the combined repertoire size of the adult and infant (the intersection). Repertoire overlap (R.O.) was calculated according to the following formula:

$$\text{R.O.} = \frac{\text{number of shared gestures}}{\text{adult's repertoire} + \text{infant's repertoire} - \text{shared gestures}}$$

All three populations contained three adult females that had successfully reproduced, so each infant's gestures were compared to those of their mother and those of two non-mother females. The repertoire overlap with the mother was higher than those of the other females in the group for 6 of the 7 infants. I used a binomial test to calculate the probability that 6 of the 7 infants would have higher overlaps with their mothers than other females given that the probability of overlap with the mother being highest was 0.33. The observed number of infants (6) whose repertoire overlap with their mothers was higher than their overlaps with other females was significantly greater than the number expected by chance (2.33) (Binomial test, $n = 7$, $p < 0.006$).

Only one infant had a higher repertoire overlap with a non-mother female than it did with its mother ($R.O._{non\text{-}mother} = 0.40$, $R.O._{mother} = 0.35$). This infant shared more gestures with its mother (14 shared) than it did with the non-mother (10 shared), but its mother had a particularly large repertoire size (32 gestures), and this caused the measure of repertoire overlap to be reduced. Figure 19 shows the comparison of each infant's repertoire overlap to the mean repertoire overlap with the other two adult females in its group.

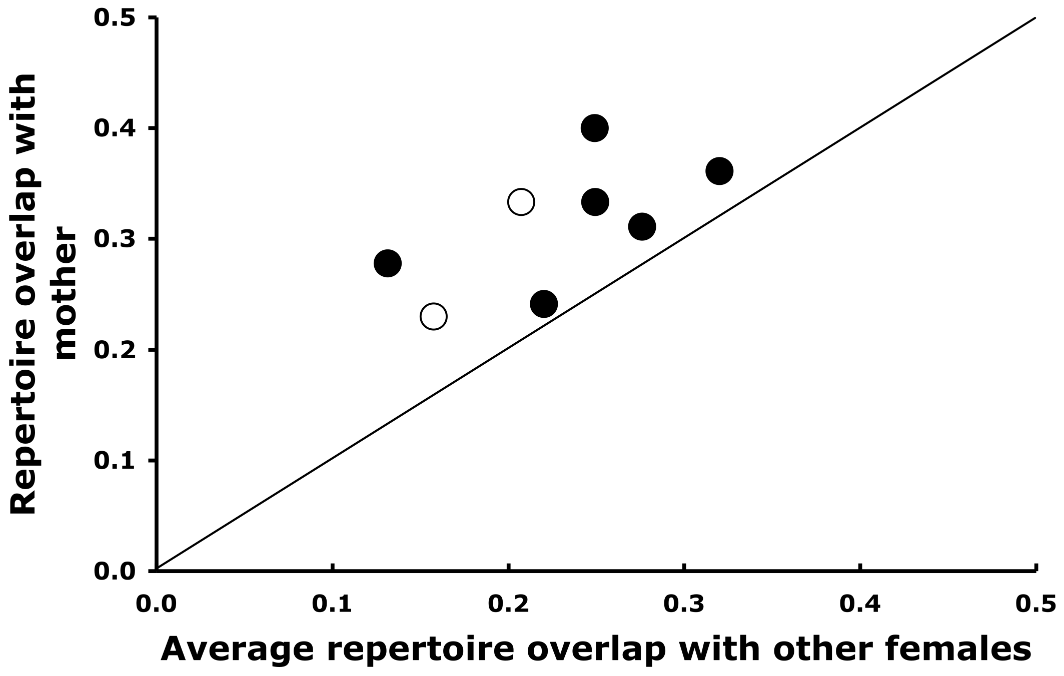


Figure 19: RATIO OF EACH INFANT'S GESTURAL REPERTOIRE OVERLAP (R.O.) WITH THEIR MOTHER TO THEIR MEAN OVERLAP WITH NON-MOTHER FEMALES.

This is plotted in the form (x,y) = (repertoire overlap with mother , average repertoire overlap with non-mother) on a Cartesian plane. The superimposed line is the predicted relationship between concordance with mother and other females if there is no difference between them. Points above the line indicate that the infant has greater concordance with the mother's gestural repertoire than it does on average with other females. Points below the line indicate would that an infant's repertoire is more similar to those of other females that to its mother. The adopted infant and juvenile are represented by open circles; their adoptive mother is considered to be their mother for this figure.

The gestural repertoires of the adopted infant and juvenile were compared to those of their adoptive and biological mothers. In both cases, the measures of repertoire overlap indicated greater similarity to the repertoire of the adoptive mother (infant/A-mother: R.O. = 0.23, juvenile/A-mother: R.O. = 0.33) than to that of the biological mother (infant/B-mother: R.O. = 0.18, juvenile/B-mother: R.O. = 0.21). Though the sample size is too small for measures of significance to be applied, comparison of the overlap values illustrates that the repertoires of the adopted individuals more closely resemble that of their adoptive or “social” mother than that of their biological mother. The ratio of the immature’s repertoire concordance with the adoptive mother to their average concordance with same-group females (including their biological mother) was compared to the same ratio in the six infants raised by their biological mothers (Figure 19). The gestural repertoires of the two adopted immatures resemble those of other females less than those of most infants raised by their biological mothers. This illustrates that adopted infants are at least as affected by their adoptive mothers as normally-reared infants are by theirs.

6.5.3 Discussion

Both age and frequency of interaction appear to play large roles in determining similarity of observed repertoires. The size of orangutan gestural repertoires is known to vary with age (Liebal 2007), with juveniles having larger repertoires than adults. This decrease of observed gestural repertoire as individuals age has also been reported in chimpanzees (Tomasello et al. 1985). Results from the analysis of gesture structure (Chapter 3) indicate that some gestures are restricted to use by certain age classes according to their differing social interactions. Tomasello et al. (1985) found chimpanzees’ use of specific gestures to be strongly associated with age class, with few gestures persisting in animals’ repertoires as they aged. The authors concluded that the low concordance between repertoires recorded at different points in time was a reflection of changes in both the infant’s social needs and its responses to different social situations as it ages. Thus, two animals that have never seen one another but are similar in age may have more similar repertoires than individuals that frequently associate with one another but are farther apart in age.

The comparison of repertoire overlap between the adopted infant and juvenile and their biological versus adoptive mother shows that closeness of association plays the primary role in determining behavioural similarity between individuals, even in the absence of any genetic relationship between individuals. Unexpectedly, the adopted immatures had lower repertoire overlaps with non-mother females than most normally-reared infants did. Though the low sample size prevents analysis of this heightened similarity to the adoptive mother, it demonstrates that the social influence of the mother figure is at least as strong in adoption as it is in related dyads, and may indicate that adopted individuals are unusually motivated to associate themselves with their adoptive mothers. To tease apart the relative contributions of association, age, and early interaction on the development and expression of a gestural repertoire, longitudinal study would be needed to trace changes in repertoire and patterns of association with other individuals during development.

6.6 Discussion of maternal influence

Infant and young juvenile orangutans are highly attuned to the actions of their mothers even when housed with many other individuals. Under captive conditions, about 5% of the day is spent with both mother and infant simultaneously directing actions towards their environment in the same way. This figure takes in only those actions done independently by each individual — excluding those actions that are necessarily simultaneous such as wrestling, holding hands, co-locomotion, or kissing. Synchronous action of mother and infant, directed ‘outwards’ at external entities, is a prime candidate for social learning. Since more than 75% of the synchronous actions observed in this study were directed towards objects or individuals external to the mother-infant dyad, the data support social learning hypotheses (van Schaik et al. 2003; e.g. Jaeggi 2006) for the development of foraging skills and social behaviour in wild orangutans.

Infant orangutans, rather than their mothers, play the major role in creating behavioural synchrony: they are more likely to initiate periods of synchrony, switching their actions to match their mothers' object-directed behaviour, but are no more likely to terminate periods of synchrony. Infants are evidently motivated to match their mother's behaviour, particularly when she is interacting with objects or other orangutans. As long-lived animals with an extended developmental period, orangutans potentially acquire valuable skills during infancy, and the mother provides the primary, if not the only, model from which they can learn. A systematic tendency to observe the mother's behaviour closely, and to attempt the same activity alongside her, would be highly adaptive for an infant orangutan, allowing skill acquisition to be guided by the social influence of the mother. I suggest that young orangutans exhibit such a tendency.

In human infants, object-directed joint attention is an important step in the developmental transition from dyadic, face-to-face interactions of infant and parent to more complex "triadic" interactions involving self, other, and object (Bakeman and Adamson 1984; Trevarthen and Aitken 2001). Triadic interaction with objects begins at 8-12 months and is believed to be important for an infant's developing understanding of attention, goals, and meaning in others as social agents (see Gómez et al. 1993). The findings of the current study indicate that infant orangutans also use the actions of the mother to guide their own object-directed behaviour. Indeed, the object-directed joint behaviour observed in orangutan mother-infant dyads may well serve much the same function that it serves in human infants, allowing the infants to gather information about both their environment and the mind of their partner – their goals and attentional state.

It is unclear what drives the motivation of the infant orangutans to copy the directed actions of their mothers. It may be that they are simply following their mothers' gaze to external referents that are particularly attractive as target objects for particular actions. Many species of primates (as well as other mammals) have shown the ability to act in response to the gaze of others (Tomasello et al. 1998; Scerif et al. 2004; Shepherd and Platt 2008). The ability to follow another's gaze to an external referent is likely a prerequisite for developing of a theory of mind. Gaze following develops in human infants between 10 and 11 months of age. It is correlated with increased language comprehension; perhaps because it allows infants to disambiguate adults' referents during language acquisition (Brooks and Meltzoff 2005). Gaze following is likely used to develop an understanding of the perceptions or intentions of others. It is not until infants are 12-18 months old that they can grasp the concept that others may see things that they themselves cannot (Moll and Tomasello 2004). Adult great apes also demonstrate this ability by following the gaze of others around barriers (Bräuer et al. 2005). There is no evidence that orangutans, or other apes, use social cues such as gaze following or behavioural synchrony to develop a complex understanding of others' understanding and beliefs about of the world. But it is possible that the orangutan infants' focus on both the actions and targets of their mothers help to construct a basic understanding of the attention and goals of others, evidenced through their selective gesture use in response to the visual attention of others and response to the directed gestures and goals of others (Chapters 3 and 4).

The effectiveness of the mother as a source of developmental guidance is shown by the fact that that the gestural repertoires of young orangutans share more gestures with those of their mothers than they do with those other females in the same social group. This finding indicates that, when gestures are socially learned, the most influential adult model is the mother. The infant's social learning of gestures could be the result of either its observation of the mother's interactions with other individuals, or its interaction with the mother as the recipient of the mother's gestures. In the latter case, described by Tomasello et al. (1985) as "second-person imitation", the infant learns to use a gesture through experience receiving it. In the wild, of course, infant orangutans will typically have no other model after whom to pattern their actions. I found that, even in captivity when surrounded by other social models, infant orangutans still focus on the behaviour of their mothers. Studies on rehabilitant orangutans have shown that infants also readily learn from humans (Russon and Galdikas 1993; Russon and Galdikas 1995), perhaps because in these circumstances humans assume the maternal role. That infant orangutans are so motivated to copy their mothers underlines the importance of the maternal relationship in development, and lends weight to arguments for vertically-transmitted culture in orangutans.

CHAPTER 7: General Discussion

The studies presented in this thesis were designed to investigate cognition and communicative abilities in orangutans through examination of the structure of their gestures and their communicative strategies. Previous studies of ape gesture primarily catalogued use of gestures and identified sensitivity to recipient's visual awareness (see Call and Tomasello 2007). Rather than focusing on the gestures themselves, I used the structure and use of orangutans' intentional gestures to investigate communicative techniques, social awareness, and theory of mind in orangutan signallers. My findings suggest that orangutans act with great sensitivity to the behaviour and awareness of others, and modify their gestures to overcome failed communicative attempts or misunderstanding.

7.1 Summary of findings

7.1.1 Orangutans use gestures intentionally

The criteria I used to determine which movements were used as intentional signals were more stringent than those used in the previous comparative studies of great ape gestures (Tomasello et al. 1989; Tanner and Byrne 1996; Pika et al. 2003; Pika et al. 2005; Liebal et al. 2006; Call and Tomasello 2007). In the previous studies, a gesture was deemed to be intentional if any individual had been observed using it in an intentional manner. My criteria required that a gesture had to have been used intentionally by an individual before it could be said that that individual used that movement as a gesture. According to this approach, all individuals in a group might exhibit a particular movement pattern but only a few might use it as an intentional gesture to communicate with others. For an individual to be said to use a movement as an intentional gesture, I had to observe the individual using the movement directed towards another individual with evidence that 1) the signaller adjusted the use or position of the gesture to the recipient's attention, 2) waited for a response, or 3) demonstrated persistence if there was no response from the recipient. Out of 71 movements initially identified as potential gestures 9 were excluded, leaving 62 actions that were deemed to be used as intentional gestures by at least some individuals. Following analysis of meaning, two new gestures were added by including extra variables in their definitions. The two new gestures, *seize* and *pull away appendage*, brought the total number of gestures to 64. Though no single orangutan used all 64 gestures, all orangutans in the study were observed using a subset of these gestures in an intentional manner.

7.1.2 Some gestures have predictable meanings

The idea of an "observable goal" was used to identify cases where it appeared that a gesture had been successful in achieving the goal of the signaller. A gesture was deemed to have an observable goal when the presumed goal of the signaller matched the outcome of the exchange. Observable goals were used in the analysis of meaning preferentially over either "presumed goal" or "outcome" in an attempt to avoid the subjectivity of the measure "presumed goal" and the possibility that the most common outcome following a particular gesture might be to reject the signaller (e.g. adults rejecting infants' requests to play). I hoped that the measure "observed goal" would be more objective than conjectures of meaning based on observer speculation, while avoiding the problem of basing meaning judgements on gestures that elicited undesired responses. Limiting the dataset to only those gestures that had observable goals reduced the number of observed examples of intentional gesture from 1421 to 698.

Slightly less than half of all 64 intentional gesture types were found to predict specific outcomes more than 70% of the time, but more than half predicted outcomes at an accuracy of 50% or more. Those gestures that predicted a single observable goal at least 70% of the time (and three times as often as the next most frequent goal) were defined as "high fidelity" meaningful gestures. Those that predicted a single observable goal at least 50% of the time (and twice as often as the next most frequent goal) were defined as "low fidelity" meaningful gestures. The gestures that failed to meet either of those thresholds, either because they were used often towards more than one goal or because they were not used at least four times towards any one goal, were labelled "unclear" gestures. Out of the 64 intentional gestures, 29 were high fidelity, 7 were low fidelity, 27 were used too infrequently to measure meaning, and 11 were used towards more than one goal. The infrequent and multi-goal gestures were grouped together as "unclear" gestures. "Affiliate/Play" was the most common observable goal (349 out of 698 tokens of gestures with observable goals were invitations to play), and had the greatest number of meaningful gestures dedicated to it (25 out of 36). I also determined that no single individual was responsible for all of the examples of secondary meaning of a gesture (Chapter 3.5). This indicated that gestures did not have completely different meanings for some individuals. A few gestures did seem to have different meanings for individuals in different age classes, but this was true across zoos and did not seem to indicate any learned meaning, but rather reflected changing social goals.

I further investigated the 11 gestures that were used towards more than one goal, in the hope of identifying variables that would define the gestures more narrowly and reduce ambiguity. Using these additional structural and social variables, I compared which variables could be used to define tokens of a gesture that would meet one of the thresholds (i.e. high-fidelity or low-fidelity) for meaning. To compare which variables were the best predictors of and accounted for the greatest amount of variance in meaning, I developed an index of effectiveness. By applying this index, I determined that the best structural predictor was the target location of the gesture (i.e. the place on the recipient's body where a tactile gesture made contact or towards which a visual gesture was directed). The best social predictor was the relationship between signaller and recipient. The analysis demonstrated that both structural and social variables are important in predicting meaning of particular gesture types. This shows that the meaning of a particular gesture type can be different between different pairings of individuals or can be altered by the signaller by directing the gesture towards a different part of the recipient's body. Though my dataset was not large enough to allow further investigation into whether different signallers alter the meaning of gestures predictably by directing them to different target locations, this may prove to be a profitable future study and would provide insight into whether gestures have meaningful parameters that can be purposefully altered in order to change the meaning of the gesture.

7.1.3 Both form and function of gestures seem to be the same for all orangutans.

Most of the gestures I identified were used by orangutans in at least two of the three zoos. In fact, 55 of the 64 gestures were used by at least one individual at two of the three zoos, and 35 of those were observed in all three zoos. Group specific gestures were extremely rare; I only observed four. I also observed four idiosyncratic gestures. The low frequency of group-specific gestures in this study supports previous findings that the gestural repertoires of orangutans in one group are as similar to those observed in other groups as they are to those within their group (Liebal et al. 2006; Liebal 2007). The seeming ubiquity of orangutan gestures implies that gestures are not often socially learned, as this process would most likely give rise to different forms of gesture (i.e. cultures) in different populations. As there is little meaningful variation between zoos, there is little evidence for cultural transmission of gesture. Meaning does not appear to vary between individuals, except when groups of individuals are defined by age class. The greatest variation in the forms (and meanings) of gestures was found between age classes. Twenty-three gestures were found to be limited to specific age classes and demonstrated in at least 2 of the 3 zoos.

The difference I observed in gesture use between age classes in orangutans supports previous findings reported from a longitudinal study of chimpanzees at Yerkes (Tomasello et al. 1989). In that study, Tomasello et al. (1989) recorded the gestures of a cohort of immature chimpanzees at two points, five years apart. They found extremely low overlap between repertoires of the same individuals measured five years apart. The authors attribute this change in exhibited gestures to changing social goals and postulate that as chimpanzees age, they either develop new gestures to address to their new social needs, or learn to apply their current gestures to new social contexts (e.g. a gesture used in play by an infant might be later co-opted for use in aggressive contexts as the individual ages and encounters a need for aggressive signals).

Is there any indication of how gestures are acquired?

Overall, the distribution and use of gestures displayed by the orangutans in this study does not suggest that social learning plays an important role in the development of gestures. However, I found that 10-25 month-old orangutan infants had repertoires more similar to those of their mothers than to those of other females, implying that some social learning might occur. If social learning of gestures happens only occasionally, and when it does it is through observation or interaction with the mother, what process is responsible for the development of other gestures? Call and Tomasello (2007) propose that gestures are acquired through a process of ontogenetic ritualisation; an individual process through which each individual learns to "ritualise" their functional actions into non-mechanically-effective gestures that elicit the same result. However, it seems unlikely that this process would always lead to the same gestural forms in each individual. When gestures resemble the first movement in a common action, they are, in essence, "intention movements" and would likely emerge in common forms in all individuals that exhibit the initial behaviour. But many gestures that exhibit intentional usage through response waiting and persistence do not resemble intention movements for effective actions (e.g. the gestures *headstand*, *dangle*, *wave*, *roll on back*, and all facial expressions do not resemble initial movements of actions).

Do orangutans understand the communicative goals of others?

Tomasello and colleagues explain that ritualisation of gestures would lead to a communication system in which individuals do not recognise others' communicative intentions through observation of their gestures, but rather learn to predict others' future behaviour by the first movements of their actions, leading to ritualized gestures. The authors explain:

> *If chimpanzees do, indeed, acquire their communicative gestures via ritualization, rather than imitation, this suggests that they do not routinely identify the communicative intentions of others as manifest in a communicative signal.* (Tomasello et al. 1997)

It seems unlikely that orangutans do not recognize the communicative intent in gestures and use them merely to predict future behaviour of the signaller. First of all, not all gestures are associated with a fully effective action that could come to be associated with the gesture. The gesture *nudge* may be exaggerated into the effective action "shove" if the recipient does not move following use of the gesture, but gestures such as *headstand* used to initiate play, or *roll on back* used as a mating request do not have effective actions associated with them. These non-intention-movement gestures could only be associated with direct actions through their association with intention-movement gestures and their corresponding actions in sequences produced when the signaller's goal is not met. This indirect associative learning process would be a much more complicated method of associating gestures with future actions than linking particular gestures to states of desire in other individuals would be.

A cognitive interpretation in which orangutans are capable of attributing basic desires or goals to other individuals may be a more parsimonious explanation than one that requires associative learning of others' future actions through observation of particular sequences of gestures and actions. A cognitive explanation is also more useful in that it allows testable predictions to be made about how orangutans should act in response to changes in the goals or visual attention of others (Byrne and Bates 2006). A particular problem with the associative learning interpretation to ape gesture is that the goals of some gestures are not achievable through effective actions, and some effective actions are never attempted between certain pairings of individuals. Though I did not collect enough examples of gestures produced during solicitation of mating to analyse consistency in meaning, I nevertheless observed the gestures *roll on back*, and *present genitals* in this context. These solicitation gestures were used by young adult females towards the adult male and were never associated with effective actions that could accomplish the females' goal. Similarly, certain pairings of signallers and recipients such as subordinate individuals gesturing towards dominant individuals never followed gestures with effective actions. Subordinate adult females frequently requested food sharing from dominant females or adult males using visual gestures, but were never observed to forcibly take food from them. Though dominant adults did not frequently share food with lower-ranking adults, I did observe several cases of such behaviour. In these cases, either the subordinate individual had, at one point, forcibly taken food from that dominant individual, or the dominant individual was able to infer the subordinate's desires from their gestures. Perhaps this inference was a generalisation from the gesture-future-action associated with other individuals that had requested food through direct action and slowly ritualized their requests into gestures. Again however, the more parsimonious explanation is that the dominant recipients were able to interpret the signallers' goals from their gestures and occasionally chose to fulfil them.

Ontogenetic ritualisation is most convincing as a proposed method of gestural acquisition for tactile gestures. Many tactile gestures are non-effective versions of actions that can be performed on other individuals to accomplish certain goals. However, their use in different contexts (and thus their meaning) may change as the signallers age. Investigation of additional social variables for gestures (that did not meet initial thresholds for meaning) revealed that the age of the signaller or the relationship between the signaller and recipient affected the meanings for 7 out of 11 gestures.

It seems most likely that orangutans acquire gestures through multiple processes. Some appear to be species-typical actions (e.g. *headstand*, *somersault*, *hit*, *bite*) that are used to initiate the contexts in which they most often appear (i.e. play). Others are likely ontogenetically ritualised from corresponding effective actions as recipients begin to anticipate the future actions of the signallers (e.g. *nudge* and *push* ritualised from shove, *arm raise* ritualised from hit). A few gestures may be socially learned through observation of or interaction with others (particularly the individual's mother). None of these processes seem capable for accounting for all of the observed gestures. Furthermore, the interactions observed in this study lead me to conclude that recipients are more likely to attribute goals to other individuals after observing their gestures than they are to come to associate each gesture from each individual with a future effective action that would force the recipient to behave in a particular way. Given orangutans' demonstrated sensitivity to the visual attention, actions, and apparent understanding of the recipient, it seems reasonable that recipients would be able to associate gestures with signallers' goals through processes other than ritualisation.

7.1.4 Orangutans gesture strategically

Analysis of orangutans' sequences of gestures demonstrated that orangutans are most likely to continue to gesture when their initial goal has not been met. This can occur either when the recipient responds in a way that does not match the signaller's goal or when the recipient fails to respond at all (Chapter 4.4.2). I also found that orangutans chose second gestures that matched the meaning of their first gestures more often than they chose second gestures of other meanings (Chapter 4.3.1). This demonstrates that orangutans use sequences of gestures to persist in their attempts to reach a specific goal.

Orangutans choose gestural modalities in accordance with the visual attention of the recipient. I observed that the orangutans used visual gestures more often when recipients could see them than they did when they could not. The recipient's gaze did not affect the frequency of tactile gestures. These findings supported the previous findings of Liebal et al. (2006) of orangutans' use of gestural modalities in response to the gaze of others, and similar findings in other great ape species from studies carried out by other members of the research group at the Max Planck in Leipzig (Pika et al. 2003; Liebal et al. 2004; Pika et al. 2005). These studies demonstrate that all great apes are able to choose gestural signals that best fit the visual attention of others, indicating that they take gaze into account when choosing how and where to communicate.

7.1.5 Orangutans take into account the recipient's gaze and responses

Investigation of the persistence strategies of orangutans following a "failed" visual gesture led me to conclude that orangutans take into account the visual attention of the recipient in both the present and the recent past. The types of gestures orangutans use at different points in a sequence indicate that they are sensitive to what they have already attempted and whether the recipient has perceived and rejected their request or possibly not perceived their gesture at all. Signallers used different types of signals depending on whether or not a recipient had seen the initial failed gesture, only holding gestures that had potentially been seen and repeating gestures more often when the recipient had not seen the first attempt. These findings demonstrate that orangutans are highly sensitive to what their recipients can and cannot see and also what they have and have not seen. This suggests that orangutans may have a basic understanding of the link between seeing and knowing. Further experiments investigating orangutans' use of strategic communication in response to the visual attention of others may shed light upon their understanding of the perceptions and awareness of others.

Sequences of gestures produced to conspecifics suggest that orangutans are able to remember what recipients have seen or witnessed, choosing gestures based on whether recipients have seen previously-attempted gestures. Experimental evidence shows more clearly that orangutans change their communicative strategies in response to the recipient's apparent understanding of their gestures. In the experimental paradigm, orangutans were presented with experimenters who were unresponsive at times and then appeared to understand, partially-understand, or misunderstand the orangutans' requests for food. The types of gestures used by the orangutans following receipt of one of the foods demonstrated that they differentiated between the three different conditions. The orangutans repeated gestures more often and used previously attempted gestures more when part of the desired food was given to them. When given the undesired food, they broadened their approach, using a wider range of gestures, avoiding those attempted before, and attempting each new gesture only once or twice. I remarked that the strategies they used to either narrow down and emphasise their requests when partially understood or to broaden their approach and try new things when misunderstood was similar to the strategies a human might use when playing charades. The selective broadening or narrowing and emphasising of gestures causes the signaller to transmit information about how close the recipient is to understanding the signaller's meaning or fulfilling the signaller's goal. This strategy could arise without having to be a conscious attempt to inform the recipient of how close he is to fulfilling the goal. The signaller could simply repeat signals that appeared to partially succeed and avoid those that clearly failed, choosing instead to attempt new signals in the hopes of landing on something that achieves the desired result.

Many of the gestures the orangutans exhibited in the experimental settings were not used as intentional gestures in interactions with conspecifics. This may be because the orangutans have a different set of gestures that they use to attract and direct the attention of the keepers, or because the distance between signaller and recipient that the experimental paradigm created forced the orangutans to use signals out with their normal repertoire. When requesting a food from a conspecific, an orangutan might reach or foodbeg with her lip, but would then move closer and likely escalate to tactile signals if the recipient did no respond. In the experimental paradigm, this was not possible, so the orangutans had to escalate or clarify their requests without moving. Perhaps this is why they used many more auditory gestures than they use in conspecific interactions. In the experiment, many of the subjects blew raspberries, banged on the enclosure, spat, kiss squeeked, or threw items. None of these actions were observed being used as intentional gestures directed toward conspecifics. The actions they used towards the experimenters that were also used with conspecifics tended to be gestures most often observed as invitations to play (i.e. *reach*, *wave*). In the experimental setting, these gestures were clearly not invitations to play but were aimed at obtaining the desired food from the experimenter. Rather than assuming that these gestures had a different meaning in the experimental context, I would posit that the gestures have a broad meaning of gaining the attention and a positive interaction from the recipient. *Reaching* towards another is a clear way of indicating one's intention to interact with that individual. *Waving* might be a good way to attract the attention of another while still communicating friendliness. Again, the artificially-imposed distance between signaller and recipient may have forced the orangutans to use signals that they would not use in natural interactions in order to attempt to elicit a positive reaction from the experimenter.

Taken together, the strategies that orangutans use in both natural and experimental interactions demonstrate that orangutans are sensitive to many aspects of potential recipients' behaviour and attention. Further study is needed to adequately assess whether orangutans learn through association to adjust their behaviour in different situations to increase their chance of receiving the desired response, or are acting in response to their projections of the perceptions and mental states of others. The most parsimonious explanation is that orangutans have some degree of a theory of mind, and are capable of attributing some perceptions and goals to other individuals. This would allow orangutans to treat others as autonomous agents, and would mean that recipients probably act in response to the desires of the signallers rather than using the signaller's gestures to merely predict the future behaviour of the signaller as has been claimed by other authors (Tomasello et al. 1997).

7.2 What abilities were likely present in our last common ancestor?

According to estimates of molecular divergence drawn from DNA comparisons, ancestral orangutans diverged from the rest of the great ape line between 10 and 13 million years ago (Stauffer et al. 2001). Though the orangutan lineage has continued to evolve, extant orangutans represent the best model for our last common ancestor, and cognitive and communicative abilities that are shared by orangutans and humans likely represent homologous traits that were present in that ancestor. Many of the findings presented in this thesis have not yet been demonstrated in other great apes, so further comparative research is necessary to ensure that the cognitive and communicative abilities I observed in orangutans are also found in the African great apes, which are more closely related to humans.

Provided similar results can be documented in African great apes, the communicative strategies and mother-infant interactions of orangutans provide insight into the social understanding of primates more than 13 million years ago. If orangutan-human similarities can be interpreted as homologous traits, it is possible to make claims about the ancestral states of infant behaviour, gestural communication, and social awareness. Based on my observations of orangutan gestural communication and mother-infant dyads as well as published reports of wild orangutans, I would make the following predictions about these types of interactions in our last common ancestor:

1) Infants would have often watched their mothers' behaviour, particularly during complex actions (Jaeggi 2006). Infants would also have been motivated to copy their mothers' actions or drawn to interact with the same item simultaneously (Chapter 6.5). This matching of the mother's directed behaviour may have been the foundation for the development of shared attention during triangular interactions that, in human children, likely helps to develop theory of mind (see Gómez et al. 1993; Trevarthen and Aitken 2001).

2) The last common ancestor would have used gestures to strategically communicate desires to other individuals based on the gaze and behaviour of others. Signallers would have used some movements as intentional gestures, and recipients would have been able to correctly interpret the signallers' goals from many of their gestures.

3) Some gestures would have predicted specific goals of the signaller, and recipients could have used the forms of the gestures along with the social context and identity of the signaller to interpret the signaller's meaning.

4) Signallers would have been able to choose the most appropriate gestural modality in relation to the gaze of the recipient, and would have been able to act in response to the gaze of the recipient in both the present and the immediate past.

5) The last common ancestor would likely have possessed a rudimentary theory of mind that is shared by extant great apes. Such an ability would allow individuals to predict the behaviour of others in different circumstances and might have incorporated the ability to take the perspective of others or at least to act in accordance with their gaze.

These speculations are based on the demonstrated communicative and behavioural flexibilities I observed in captive orangutans. Further testing is necessary to clarify the nature of social understanding in captive orangutans. Captive groups of orangutans provide a unique opportunity to observe complicated social interactions in orangutans. As wild-living orangutans are fairly solitary aside from mother-offspring pairings (van Noordwijk and van Schaik 2005), they do not have as many opportunities to interact with other individuals in complex social situations as captive orangutans who have been housed in social groups. It is possible that the abilities I observed in the gestures between group members (particularly unrelated individuals) emerged only when the animals were forced into group-living situations. Observational evidence is needed from populations in the wild to determine whether orangutans exhibit any of these social or cognitive abilities in their natural interactions or communication.

7.3 Can ape gesture tell us anything about the evolution of cognition and language?

Researchers interested in learning about the evolutionary origins of language have focused their studies mainly on the vocal communication of non-human primates. These studies of vocal communication have found particularly exciting results in the area of alarm calls, which appear to serve as semantic signals in many species. Alarm calls are an ideal type of signal to use for playback studies in which a recorded call is broadcast to a group of animals whose responses are then recorded, as they tend to elicit either predator evasion responses or calling behaviour. Many species of monkey have been determined to have functionally-referential alarm calls through studies of this design (Cheney and Seyfarth 1980; Zuberbühler 2000; Zuberbühler 2001; Zuberbühler 2003; Kirchhof and Hammerschmidt 2006). The term "functionally-referential" refers to the difficulty in determining whether information is purposefully transmitted—calls transmit information about external referents, but no assumption is made about whether the signaller intends to inform the recipient of something through the vocalization (Zuberbühler 2003). Alarm calls may be the most likely to produce responses in the listening animals, but because they are extremely important signals to the survival of the individuals within the group, there is pressure that they be "honest signals" and be difficult to modify or fake (Knight 1998). Perhaps unsurprisingly, the structure of alarm calls appears to be, to a large extent, "fixed" and unable to be altered (Corballis 2002), though the context in which the calls are given displays greater flexibility (Seyfarth and Cheney 1986). Alarm calls are, perhaps, the least likely type of vocalisation to show flexibility in form or use, as the consequences of misinterpreting a call given by another could be deadly. Out with the realm of alarm calls, there is only slight evidence of referential labelling in ape vocalisations (see Slocombe and Zuberbühler 2005).

Primates do not seem capable of vocal learning on a scale comparable to humans, birds, or cetaceans, but there is some indication of limited flexibility in the structure or the deployment of calls (Slocombe and Zuberbühler 2005; Slocombe and Zuberbühler 2007). There is also evidence that primate calls may be shaped to some degree to fit in with the calls of others either in their form (Marshall et al. 1999) or in their use (Seyfarth and Cheney 1986).

Though some flexibility has been observed within primate vocal communication systems, primates' vocalisations nonetheless display less flexibility than their gestures (Tomasello and Call 1997; Corballis 2002; Call and Tomasello 2007). Auditory signals have the obvious advantage that they can be used to communicate with unseen individuals at a distance. So-called "auditory gestures" may allow primates to produce audible signals with greater flexibility than is possible in their vocal systems. African apes extend their gestures into the auditory modality be exhibiting a large range of auditory gestures (see Call and Tomasello 2007). Rhythmic auditory behaviours have been observed in both *Pan* (drumming) (Goodall 1986; Arcadi et al. 1998; Arcadi et al. 2004) and *Gorilla* (chest beating) (Schaller 1963), but never in orangutans (Fitch 2006). Orangutans' gestures are only rarely auditory (see Chapter 3) and on only two occasions did I observe an orangutan accompanying a gesture with a vocalisation. Chimpanzees, on the other hand, not only use both gestural and vocal signals, but also sometimes combine the two, producing synchronised multi-modal signals (Hopkins, pers. comm.).

Despite the prevalence of drumming and chest-beating in the African apes, the majority of great ape gestures are of either visual or tactile modalities. The orangutans in this study used 39 visual and 25 tactile gestures (see Chapter 3.4). Simone Pika and Katja Liebal identified between 20 and 40 total gestures for each ape species in their cross-species comparison of repertoire size (Pika et al. 2003; Pika et al. 2005; Liebal et al. 2006). Though they found some differences in the relative numbers of visual to tactile gestures, they found no difference in gestural repertoire size amongst the great apes.

None of the studies of ape gesture thus far have discovered evidence that gestures are learned or culturally transmitted. This is of particular relevance in attempting to relate ape gesture to the evolution of language. Following the initial failure to teach a chimpanzee to speak (Hayes and Hayes 1951) and the relative success of teaching several apes to sign (Gardner and Gardner 1969; Patterson 1978; Miles 1990), scientists began to speculate that perhaps language evolved first in the gestural rather than vocal modality. There have been several versions of the "gesture first" theory of language evolution (Hewes 1973; Rizzolatti and Arbib 1998; Corballis 2002; Volterra et al. 2005; Armstrong and Wilcox 2007), but all of them require a transition during which speech takes over from gesture as the dominant modality. There is no evidence that such a switch occurred, and from the evidence of modern signed languages, also no apparent necessity. It is difficult then to construct a scenario in which language developed fully (or to an advanced degree) in one modality and then switched to an entirely different modality. David McNeill argues that language evolved in both modalities simultaneously, each serving a different purpose but coordinated in their expression (McNeill 2007). As the orangutans I observed vocalised only rarely and never in conjunction with a gesture, the current study can shed no light on the communicative potential of a gestural versus a multimodal system. However, the increase in vocal behaviour and auditory gestures in the African great apes and the purported coordination of vocalisation and gesture in chimpanzees may support a multimodal view of language evolution as African apes (particularly of the genus *Pan*) share a longer evolutionary history with humans. Perhaps the communicative abilities that orangutans express through gesture are expressed in both vocal and gestural modalities in chimpanzees and bonobos. If this is the case, then the distribution of gesture types in the extant great apes may inform us about the progression of communicative modalities in the evolution of language. However, it is also possible that the social systems in the different apes (see McGrew et al. 1996) could also account for the varied distribution of gesture types.

Orangutans appear to act as if they are sensitive to the perceptions and desires of others. They use their gestures in flexible ways: substituting gestures with similar meanings to persist towards a goal, repeating gestures that may not have been seen by the recipient, and using a wide range of untried gestures if they have been “misunderstood.” The forms of their gestures, however, are relatively consistent from one population to the next. The gestures of the other ape species seem to exhibit great similarity between populations as well (see Call and Tomasello 2007). It is clear from sign language studies and other training programs that apes can learn to use novel movements as meaningful signs or gestures to communicate with others. However, they do not seem to develop the regional cultures of gestures one would expect if apes copied gestures from one another. There have only been a very few “group-specific” gestures observed in great apes. The only gestures I observed that were limited to one population were three facial expressions (Chapter 3.4.3) and the gesture/behaviour *tandem walk* (which was observed in all zoos but only used intentionally in one). All four gestures were only observed in two individuals. Liebal (2006) documented a unique food sharing gesture in one group of orangutans, but again, it was used by only a few individuals. The greatest variation in ape gestures appears to be on the individual level. An individual may have several gestures that he or she does not share with any other members of the group or even of the species. A single male gorilla at the San Francisco zoo was reported to use several gestures that were not used by other members of his group and have not been reported in other gorilla groups; notably iconic and deictic gestures (Tanner and Byrne 1996). Though some have disputed the idea iconicity of his gestures (Tomasello and Call 2007), they were certainly unique within his group.

Variation in gestural form seems to take place mainly on an individual rather than group level. While these observations do not indicate cultural transmission of gestures, they nonetheless demonstrate that forms are flexible and individuals can vary in their form or use of gestures. Ape gestures do not appear to vary much in form, but are used in highly flexible combinations in response to the subtleties of the social environment. I believe that it is the flexibility in deployment rather than structure that makes gesture an ideal medium for investigating cognition in the signaller, and may be a qualitative difference distinguishing gesture from vocalisation in primates. The strategies that I observed in orangutans suggest that signallers are sensitive to the perceptions and reactions of their intended recipients. Though the gestures themselves do not vary to any great degree, the sequences in which they are deployed and the purposeful deployment of different modalities coupled with waiting for a response, demonstrate that orangutans have specific desires and act flexibly and strategically to elicit certain behaviour from others. If we take the strategies present in orangutan communication to represent ancestral states of hominin communicative ability, it seems likely that the understanding of others as intentional agents and the ability to act in accordance with the perceptions of others predate any significant increase in communicative complexity.

Orangutans appear to possess a fairly sophisticated ability to respond to others' perceptions of and reactions to their communicative gestures, but do not alter the structure of the gestures themselves to any great degree. Within their limited range of signals, orangutans are able to communicate specific meanings, reiterate those meanings, and transmit information about how well a recipient has understood their desires. The gestural communication of orangutans demonstrates how an animal can communicate flexibly and with great sensitivity to the actions and perceptions of others within a system of non-culturally-transmitted signals. By selectively repeating gestures or choosing new ones, orangutans may help recipients to arrive at shared meanings more quickly. By selectively matching to object-directed actions of their mothers, orangutan infants engage in triangular interactions and may begin to share their mothers' attention and develop an understanding of perceptions in others. The communicative strategies, sensitivity to the gaze of others, and participation in triadic interactions demonstrated by orangutans provide a behavioural foundation for the emergence of sophisticated social understanding later in the evolution of hominin cognition.

References

Aiello, L. C., Runciman, W. G., Maynard-Smith, J. and Dunbar, R. I. M. (1996). Terrestriality, bipedalism and the origin of language. *Proceedings of the British Academy* 88: 269-289.

Alibali, M. W. and Goldin-Meadow, S. (1993). Gesture speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology* 25(4): 468-523.

Alibali, M. W. and Goldin-Meadow, S. (1993). Gesture-Speech Mismatch and Mechanisms of Learning: What the Hands Reveal about a Child's State of Mind. *Cognitive Psychology* 25(4): 468-523.

Ambrose, S. H. (2001). Paleolithic technology and human evolution. *Science* 291: 1748-1753.

Arbib, M. A. (2002). The mirror system, imitation, and the evolution of language. Imitation in Animals and Artifacts. Complex Adaptive Systems. K. Dautenhahn and C. L. Nehaniv. Cambridge, Masachusetts, USA, MIT Press: 229-280.

Arbib, M. A. (2005). From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences* 28: 105-167.

Arcadi, A. C., Robert, D. and Boesch, C. (1998). Buttress drumming by wild chimpanzees: Temporal patterning, phrase integration into loud calls, and preliminary evidence for individual distinctiveness. *Primates* 39(4): 505-518.

Arcadi, A. C., Robert, D. and Mugurusi, F. (2004). A comparison of buttress drumming by male chimpanzees from two populations. *Primates* 45(2): 135-139.

Armstrong, D. F., Stokoe, W. C. and Wilcox, S. E. (1995). Gesture and the nature of language. Cambridge, Cambridge University Press.

Armstrong, D. F. and Wilcox, S. E. (2007). The gestural origin of language. Oxford, Oxford University Press.

Arnold, K. and Zuberbuhler, K. (2006). Semantic combinations in primate calls. *Nature* 441(7091): 303-303.

Bakeman, R. and Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development* 55(4): 1278-1289.

Bard, K. A. (1992). Intentional Behaviour and Intentional Communication in Young Free-Ranging Orangutans. *Child Development* 63: 1186-1197.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L. and Volterra, V. (1979). The Emergence of Symbols. New York, Academic Press.

Bekoff, M. and Byers, J. A., Eds. (1998). Animal play: evolutionary, comarative and ecological perspectives. Cambridge, Cambridge University Press.

Bickerton, D. (1990). Language and species. Chicago, Chicago Univ. Press.

Biro, D., Inoue-Nakamura, N., Tonooka, R., Yamakoshi, G., Sousa, C. and Matsuzawa, T. (2003). Cultural innovation and transmission of tool use in

wild chimpanzees: evidence from field experiments. *Animal Cognition* 6(4): 213-223.

Botha, R. (2007). On homesign systems as a potential window on language evolution. *Language & Communication* 27(1): 41-53.

Boysen, S. T., Kuhlmeier, V. A., Halliday, P. and Halliday, Y. M. (1999). Tool use in captive gorillas. The mentalities of gorillas and orangutans: Comparative perspectives. S. T. Parker, R. W. Mitchell and H. L. Miles. Cambridge, Cambridge University Press: 179-187.

Brain, C. K. and Sillen, A. (1988). Evidence from the Swartkrans cave for the earliest use of fire. *Nature* 336: 464-466.

Bräuer, J., Call, J. and Tomasello, M. (2005). All Great Ape species follow gaze to distant locations and around barriers. *Journal of Comparative Psychology* 119(2): 145-154.

Brooks, R. and Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science* 8(6): 535-543.

Bruner, J. (1981). Intention in the structure of action and interaction. Advances in Infancy Research. L. Lipsitt. New Jersey, Ablex, Norwood. 1: 41-56.

Byrne, R. W. (1995). The thinking ape: evolutionary origins of intelligence. Oxford, Oxford Univ. Press.

Byrne, R. W. and Bates, L. A. (2006). Why are animals cognitive? *Current Biology* 16(12): 445-448.

Byrne, R. W. and Whiten, A. (1988). Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans. Oxford, Clarendon Press.

Byrne, R. W. and Whiten, A. (1992). Cognitive Evolution in Primates - Evidence from Tactical Deception. *Man* 27(3): 609-627.

Call, J. (2003). Beyond learning fixed rules and social cues: abstraction in the social arena. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 358(1435): 1189-1196.

Call, J., Hare, B., Carpenter, M. and Tomasello, M. (2004). 'Unwilling' versus 'unable': chimpanzees' understanding of human intentional action. *Developmental Science* 7(4): 488-498.

Call, J. and Tomasello, M. (1994). Production and Comprehension of Referential Pointing by Orangutans (Pongo-Pygmaeus). *Journal of Comparative Psychology* 108(4): 307-317.

Call, J. and Tomasello, M. (2007). The gestural communication of apes and monkeys. Mahwah, NJ, Lawrence Erlbaum Associates, Inc.

Call, J. and Tomasello, M. (2007). The gestural repertoire of chimpanzees (*Pan troglodytes*). The gestural communication of apes and monkeys. J. Call and M. Tomasello. Mahwah, NJ, Lawrence Erlbaum Associates, Inc.

Cartmill, M. and Smith, F. (in press). The Human Lineage (Foundation of Human Biology). Wilmington, DE, Wiley-Liss.

Cavalli-Sforza, L. L., Piazza, A., Menozzi, P. and Mountain, J. (1988). Reconstruction of human evolution: bringing together genetic, archeological, and linguistic data. *Procedings of the National Academy of Science* 85: 6001-6006.

Cheney, D. L. and Seyfarth, R. M. (1980). Vocal Recognition in Free-Ranging Vervet Monkeys. *Animal Behaviour* 28(MAY): 362-&.

Cheney, D. L. and Seyfarth, R. M. (1990). How Monkeys See the World. Chicago and London, University of Chicago Press.

Chomsky, N. (1957). Syntactic structures. The Hague, Mouton.

Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge, Mass, MIT Press.

Chomsky, N. (1972). Language and mind. New York, Harcourt, Brace & Javanovich.

Church, R. B. and Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition* 23: 43-71.

Corballis, M. C. (2002). From Hand to Mouth, the Origins of Language. Princeton, New Jersey, Princeton University Press.

d'Errico, F., Henshilwood, C., Vanhaeren, M. and van Niekerk, K. (2005). Nassarius kraussianus shell beads from Blombos Cave: evidence for symbolic behaviour in the Middle Stone Age. *Journal of Human Evolution* 48(1): 3-24.

Darwin, C. (1955). The Expression of the Emotions in Man and Animals. New York, Philosophical Library.

Dickinson, A. and Balleine, B. (1994). Motivational control of goal-directed action. *Learning and behavior* 22: 1-18.

Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S. L., Wiebe, V., Kitano, T., Monaco, A. P. and Paebo, S. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 418: 869-872.

Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences* 4(7): 258-267.

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition* 100(1): 173-215.

Fouts, R. S., Fouts, D. H. and van Cantfort, T. E. (1989). The infant Loulis learns signs from cross-fostered chimpanzees. Teaching sign language to chimpanzees. R. A. Gardner, B. T. Gardner and T. E. Cantfort. Albany, State University of New York Press: 280-292.

Fouts, R. S. and Mills, S. (1997). Next of kin. London, Michael Josph.

Fox, E. A., Sitompul, A. F. and van Schaik, C. P. (1999). Intelligent tool use in wild Sumatran orangutans. The Mentalities of Gorillas and Orangutans. S. T. Parker, R. W. Mitchell and H. L. Miles. Cambridge, University Press.

Galdikas, B. M. F. (1983). The orangutan long call and snag crashing at Tanjung Puting Reserve. *Primates* 24(3): 371-384.

Galdikas, B. M. F. and Wood, J. W. (1990). Birth Spacing Patterns in Humans and Apes. *American Journal of Physical Anthropology* 83(2): 185-191.

Gannon, P. J., Holloway, R. L., Broadfield, D. C. and Braun, A. R. (1998). Asymmetry of chimpanzee planum temporale: humanlike pattern of Wernicke's brain language area homolog. *Science* 279: 220-222.

Gardner, R. A. and Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science* 165: 664-672.

Geissmann, T. (1999). Duet songs of the siamang, Hylobates syndactylus: II. Testing the pair-bonding hypothesis during a partner exchange. *Behaviour* 136(8): 1005-1039.

Gershkoff-Stowe, L. and Goldin-Medow, S. (2002). Is there a natural order for expressing semantic relations? *Cognitive Psychology* 45(3): 375-412.

Goldin-Meadow, S. (2003). Hearing Gesture: How Our Hands Help Us Think. Cambridge, Mass, The Belknap Press of Harvard University Press.

Goldin-Meadow, S. (2005). What language creation in the manual modality tells us about the foundations of language. *Linguistic Review* 22(2-4): 199-225.

Goldin-Meadow, S., Alibali, M. W. and Church, R. B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review* 100(2): 279-297.

Goldin-Meadow, S., Gelman, S. A. and Mylander, C. (2005). Expressing generic concepts with and without a language model. *Cognition* 96(2): 109-126.

Goldin-Meadow, S., So, W. C., Özyürek, A. and Mylander, C. (in press). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences of the United States of America*.

Golinkoff, R. M. (1986). ‘‘I beg your pardon?’’: The pre-verbal negotiation of failed messages. *Journal of Child Language* 20: 199-208.

Gómez, J. C. (1998). Assessing theory of mind with nonverbal procedures: Problems with training methods and an alternative "key" procedure. *Behavioral and Brain Sciences* 21(1): 119-+.

Gómez, J. C. (2004). Apes, monkeys, children, and the growth of mind. Cambridge, Massachusetts, and London, England, Harvard University Press.

Gómez, J. C., Sarriá, E. and Tamarit, J. (1993). The comparative study of early communication and theories of mind: ontogeny, phylogeny, and pathology. Understanding other minds. S. Baron-Cohen, H. Tager-Flusberg and D. Cohen. Oxford, Oxford University Press: 397-426.

Goodall, J. (1986). The Chimpanzees of Gombe: Patterns of Behavior. Cambridge Massachusetts, Belknap Press of Harvard University Press.

Greenfield, P. M. and Savage-Rumbaugh, E. S. (1990). Grammatical combinations in Pan paniscus: Processes of learning and invention in the evolution and development of language. “Language” and intelligence in monkeys and apes: comparative developmental perspectives. S. T. Parker and K. R. Gibson. New York, Cambridge University Press: 540-578.

Griffin, E. (1997). A first look at communication theory. New York, McGraw Hill.

Hare, B., Call, J., Agnetta, B. and Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour* 59(4): 771-785.

Hare, B. and Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour* 68: 571-581.

Harvey, P. H. and Clutton-Brock, T. H. (1985). Life history variation in primates. *Evolution* 39(3): 559-581.

Hauser, M. D., Chomsky, N. and Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science* 298(5598): 1569-1579.

Hauser, M. D., Teixidor, P., Field, L. and Flaherty, R. (1993). Food-elicited calls in chimpanzees: Effects of food quantity and divisibility? *Animal Behaviour* 45: 817-819.

Hauser, M. D. and Wrangham, R. W. (1987). Manipulations of food calls in captive chimpanzees: A preliminary report. *Folia Primatologica* 48: 207-210.

Hayes, C. (1951). The ape in our house. New York, Harper.

Hayes, K. J. and Hayes, C. (1951). The intellectual development of a home-raised chimpanzee. *Proceedings of the American Philosophical Society* 95: 105-109.

Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Current Anthropology* 12(1-2): 5-24.

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences* 21(1): 101-+.

Hinde, R. A. and Spencer-Booth, Y. (1967). The behaviour of socially living rhesus monkeys in their first two and a half years. *Animal Behaviour* 15(1): 169-196.

Hirata, S. and Celli, M. L. (2003). Role of mothers in the acquisition of tool-use behaviours by captive infant chimpanzees. *Animal Cognition* 6(4): 235-244.

Humboldt, W. (1836/1999). On Language; on the diversity of human language construction and its influence of the mental development of the human species. Cambridge, Cambridge University Press.

Hunt, G. R. (1996). Manufacture and use of hook-tools by New Caledonian crows. *Nature* 379(6562): 249-251.

Hunt, G. R. (2000). Human-like, population-level specialization in the manufacture of pandanus tools by New Caledonian crows *Corvus moneduloides*. *Proc. R. Soc. London* 267: 403-413.

Ingmanson, E. J. (1996). Tool-using behavior in wild Pan paniscus: Social and ecological considerations. Reaching into thought: The minds of the great apes. A. E. Russon and K. A. Bard. Cambridge, Cambridge University Press: 190-210.

Inoue-Nakamura, N. and Matsuzawa, T. (1997). Development of stone tool use by wild chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 111(2): 159-173.

Jackendoff, R. (2002). Foundations of language: brain, meaning, grammar, evolution, OUP.

Jaeggi, A. (2006). The role of social learning in the acquisition of foraging skills in wild Bornean orang-utans (*Pongo pygmaeus*); The roots of culture. Anthropological Institute and Museum. Zürich, University of Zürich. Diploma: 46.

Janik, V. M. and Slater, P. J. B. (1997). Vocal learning in mammals. *Advances in the study of behavior* 26: 59-99.

Jensvold, M. and Fouts, R. (1993). Imaginary Play in chimpanzees (Pan troglodytes). *Human Evolution* 8(3): 217-227.

Kako, E. (1999). Elements of syntax in the systems of three language-trained animals. *Animal Learning & Behavior* 27(1): 1-14.

Karakashian, S. J., Gyger, M. and Marler, P. (1988). Audience effects on alarm calling in chickens (Gallus gallus). *Journal of Comparative Psychology* 102: 129-135.

Kendon, A. (1991). Some considerations for a theory of language origins. *Man* 26(2): 199-221.

King, B. J. (2004). The dynamic dance; Nonvocal communication in African great apes. Cambridge, Massachusetts, Harvard University Press.

King, M. C. and Wilson, A. C. (1975). Evolution at two levels in humans and chimpanzees. *Science* 188: 107-116.

Kirby, S., Dowman, M. and Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences of the United States of America* 104(12): 5241-5245.

Kirchhof, J. and Hammerschmidt, K. (2006). Functionally referential alarm calls in tamarins (Saguinus fuscicollis and Saguinus mystax) - evidence from playback experiments. *Ethology* 112(4): 346-354.

Knight, C. (1998). Ritual/speech co-evolution: a solution to the problem of deception. Approaches to the Evolution of Language: Social and Cognitive Bases. J. R. Hurford, M. Studdert-Kennedy and C. Knight. Cambridge, UK, Cambridge University Press: 68-91.

Krause, J., Lalueza-Fox, C., Orlando, L., Enard, W., Green, R. E., Burbano, H. A., Hublin, J.-J., Hanni, C., Fortea, J., de la Rasilla, M., Bertranpetit, J., Rosas, A. and Paabo, S. (2007). The derived FOXP2 variant of modern humans was shared with Neandertals. *Current Biology* 17(21): 1908-1912.

Krutzen, M., Mann, J., Heithaus, M. R., Connor, R. C., Bejder, L. and Sherwin, W. B. (2005). Cultural transmission of tool use in bottlenose dolphins. *Proceedings of the National Academy of Sciences of the United States of America* 102(25): 8939-8943.

Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. and Monaco, A. P. (2001). A novel transcription factor is mutated in a severe speech and language disorder. *American Journal of Medical Genetics* 105(7): O80.

Landis, J. R. and Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33(1): 159-174.

Leakey, M. D. and Hay, R. L. (1979). Pliocene footprints in the Laetolil Beds at Laetoli, northern Tanzania. *Nature* 278: 317-323.

Leavens, D. A. and Hopkins, W. D. (1999). The whole-hand point: The structure and function of pointing from a comparative perspective. *Journal of Comparative Psychology* 113(4): 417-425.

Leavens, D. A., Hostetter, A. B., Wesley, M. J. and Hopkins, W. D. (2004). Tactical use of unimodal and bimodal communication by chimpanzees, *Pan troglodytes*. *Animal Behaviour* 67: 467-476.

Leavens, D. A., Russell, J. L. and Hopkins, W. D. (2005). Intentionality as measured in the persistence and elaboration of communication by chimpanzees (Pan troglodytes). *Child Development* 76: 291-376.

Liebal, K. (2007). The gestural communication of orangutans. The gestural communication of monkeys and apes. J. Call and M. Tomasello. Mahwah, NJ, Lawrence Erlbaum Associates, Inc.

Liebal, K., Call, J. and Tomasello, M. (2004). Use of gesture sequences in chimpanzees. *American Journal of Primatology* 64(4): 377-396.

Liebal, K., Pika, S., Call, J. and Tomasello, M. (2004). Great ape communicators move in front of recipients before producing visual gestures. *Interaction studies* 5(2): 199-219.

Liebal, K., Pika, S. and Tomasello, M. (2004). Social communication in siamangs (Symphalangus syndactylus): use of gestures and facial expressions. *Primates* 45(1): 41-57.

Liebal, K., Pika, S. and Tomasello, M. (2006). Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture* 6(1): 1-38.

Liszkowski, U., Carpenter, M., Henning, A., Striano, T. and Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science* 7(3): 297-307.

Lock, A. (2001). Preverbal communication. Blackwell Handbook of Infant Development. J. G. Bremner and A. Fogel. Oxford, Oxford University Press: 379–403.

Lonsdorf, E. (2006). What is the role of mothers in the acquisition of termite-fishing behaviors in wild chimpanzees (*Pan troglodytes schweinfurthii*)? *Animal Cognition* 9: 36-46.

MacKinnon, J. (1974). The behaviour and ecology of wild orang-utans (Pongo pygmaeus). *Animal Behaviour* 22: 3-74.

Marler, P., Dufty, A. and Pickert, R. (1986). Vocal communication in the domestic chicken: I. Does a sender communicate information about the quality of a food referent to a receiver? *Animal Behaviour* 34: 188-193.

Marshall, A. J., Wrangham, R. W. and Arcadi, A. C. (1999). Does learning affect the structure of vocalizations in chimpanzees? *Animal Behaviour* 58: 825-830.

Matheson, M. D., Cooper, M., Weeks, J., Thompson, R. and Fragaszy, D. (1998). Attribution is more likely to be demonstrated in more natural contexts. *Behavioral and Brain Sciences* 21(01): 124-126.

McFarland, D. and Bösser, T. (1993). Intelligent behavior in animals and robots. Cambridge, MA, MIT Press.

McGrew, W. C., Ham, R. M., White, L. J. T., Tutin, C. E. G. and Fernandez, M. (1997). Why don't chimpanzees in Gabon crack nuts? *International Journal of Primatology* 18(3): 353-374.

McGrew, W. C., Marchant, L. F. and Nishida, T. (1996). Great Ape Societies. Cambridge, Cambridge University Press.

McNeill, D. (1992). Hand and Mind. Chicago, The University of Chicago Press.

McNeill, D. (2000). Language and Gesture. Cambridge, Cambridge University Press.

McNeill, D. (2007). Gesture and Thought. Chicago, University of Chicago Press.

Miles, H. L. and Roberts, W. P. (1998). Methodologies, not method, for primate theory of mind. *Behavioral and Brain Sciences* 21(01): 126-127.

Miles, H. L. W. (1990). The cognitive foundations for reference in a signing orangutan. "Language" and intelligence in monkeys and apes. Comparative developmental perspectives. S. T. Parker and K. R. Gibson. Cambridge, Cambridge Univ. Press.

Mitani, J. C. (1985). Mating-Behavior of Male Orangutans in the Kutai Game Reserve, Indonesia. *Animal Behaviour* 33(MAY): 392-402.

Moll, H. and Tomasello, M. (2004). 12-and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science* 7(1): F1-F9.

Montagu, A. (1976). Toolmaking, hunting, and the origin of language. Origins and evolution of language and speech. S. R. Harnad, H. D. Steklis and J. Lancaster. New York, The New York Academy of Sciences. 280: 266-274.

Morford, J. P. and Goldin-Meadow, S. (1997). From here and now to there and then: The development of displaced reference in homesign and English. *Child Development* 68(3): 420-435.

Nobe, S. (2000). Where do most spontaneous representational gestures actually occur with respect to speech? Language and Gesture. D. McNeill. Cambridge, Cambridge University Press: 186-198.

O'Grady, W., Dobrovolsky, M. and Aronoff, M. (1997). Contemporary Linguistics. Boston, Bedford/St. Martin's.

Ottoni, E. B. and Mannu, M. (2001). Semifree-ranging tufted capuchins (Cebus apella) spontaneously use tools to crack open nuts. *International Journal of Primatology* 22(3): 347-358.

Parker, S. T., Kerr, M., Markowitz, H. and Gould, J. (1999). A survey of tool use in zoo gorillas. The mentalities of gorillas and orangutans: Comparative perspectives. S. T. Parker, R. W. Mitchell and H. L. Miles. Cambridge, Cambridge University Press: 188-193.

Patterson, F. (1978). Linguistic capabilities of a lowland gorilla. Sign Language and Language Acquisition in Man and Ape. F. C. C. Peng. Boulder, CO, Westview Press: 161-201.

Patterson, F. and Linden, E. (1981). The education of Koko. New York, Owl Books.

Patterson, F. G. (1980). Innovative uses of language by a gorilla: a case study. Children's Language. K. Nelson. New York, Gardner. vol. 2: 497-561.

Patterson, F. G. and Cohn, R. H. (1990). Language acquisition by a lowland gorilla: Koko's first ten years of vocabulary development. *Word* 41(2): 97-143.

Piaget, J. (1952). The Origins of Intelligence in Children. New York, Norton.

Pika, S. (in press). What is the nature of the gestural communication of great apes? The Shared Mind. J. Zlatev, T. Racine, C. Sinha and E. Itkonen.

Pika, S., Liebal, K. and Tomasello, M. (2003). Gestural communication in young gorillas (*Gorilla gorilla*): Gestural repertoire, learning and use. *American Journal of Primatology* 60(3): 95-111.

Pika, S., Liebal, K. and Tomasello, M. (2005). Gestural communication in subadult bonobos (*Pan paniscus*): Gestural repertoire and use. *American Journal of Primatology* 65(1): 39-61.

Pika, S., Liebal, K. and Tomasello, M. (2005). Gestural communication in subadult bonobos (Pan paniscus): Repertoire and use. *American Journal of Primatology* 65(1): 39-61.

Pollick, A. S. and de Waal, F. B. M. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences of the United States of America* 104(19): 8184-8189.

Posada, S. and Colell, M. (2007). Another gorilla (Gorilla gorilla gorilla) recognizes himself in a mirror. *American Journal of Primatology* 69(5): 576-583.

Poss, S. R., Kuhar, C., Stoinski, T. S. and Hopkins, W. D. (2006). Differential use of attentional and visual communicative signaling by orangutans (Pongo

pygmaeus) and gorillas (Gorilla gorilla) in response to the attentional status of a human. *American Journal of Primatology* 68(10): 978-992.

Povinelli, D. J., Bierschwale, D. T. and Cech, C. G. (1999). Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. *British Journal of Developmental Psychology* 17: 37-60.

Povinelli, D. J., Boysen, S. T. and Nelson, K. E. (1990). Inferences About Guessing and Knowing by Chimpanzees (Pan-Troglodytes). *Journal of Comparative Psychology* 104(3): 203-210.

Povinelli, D. J., Perilloux, H. K., Reaus, J. E. and Bierschwale, D. T. (1998). Young and juvenile chimpanzees' (*Pan troglodytes*) reactions to intentional versus accidental and inadvertent actions. *Behavioural Processes* 42: 2005-218.

Povinelli, D. J., Rulf, A. B. and Bierschwale, D. T. (1994). Absence of Knowledge Attribution and Self-Recognition in Young Chimpanzees (Pan-Troglodytes). *Journal of Comparative Psychology* 108(1): 74-80.

Quintilian. (2006). "Institutes of oratory." Retrieved December 26, 2007, from http://honeyl.public.iastate.edu/quintilian/.

Rijksen, H. D. (1978). A Field Study on Sumatran Orangutans. Wageningen, Mededelingen Landbouwhogeschool.

Rizzolatti, G. and Arbib, M. A. (1998). Language within our grasp. *Trends in Neuroscience* 21(5): 188-194.

Rizzolatti, G., Fadiga, L., Gallese, V. and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research* 3: 131-141.

Roberts, S. (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes* 7: 242-268.

Rothmann, M. and Teuber, E. (1915). Aus der Anthropoidenstation auf Teneriffa: Ziele und Aufgaben der Station sowie erste Beobachtungen an den auf ihr gehaltenen Schimpansen. *Abhandlungen der koniglich preussischen Akademie der Wissenschaften (physikalisch-mathematische Klasse)* 2: 1-20.

Roush, R. S. and Snowdon, C. T. (2000). Quality, quantity, distribution and audience effetcs on food calling in cotton-top tamarins. *Ethology* 106: 673-690.

Russon, A. E. (1996). Imitation in everyday use: Matching and rehearsal in the spontaneous imitation of rehabilitant orangutans (*Pongo pygmaeus*). Reaching into thought. The minds of the great apes. A. E. Russon, K. A. Bard and S. T. Parker. Cambridge, Cambridge University Press: 152-176.

Russon, A. E. and Galdikas, B. M. F. (1993). Imitation in free-ranging rehabilitant orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology* 107(2): 147-161.

Russon, A. E. and Galdikas, B. M. F. (1995). Constraints on Great Apes Imitation - Model and Action Selectivity in Rehabilitant Orangutan (Pongo-Pygmaeus) Imitation. *Journal of Comparative Psychology* 109(1): 5-17.

Russon, A. E., Handayani, D. P., Kuncoro, P. and Ferisa, A. (2007). Orangutan leaf-carrying for nest-building: Toward unraveling cultural processes. *Animal Cognition* 10(2): 189-202.

Ruvolo, M. (1997). Molecular phylogeny of the hominopids: inferences from multiple independent DNA sequence data sets. *Molecular Biology and Evolution* 14: 248-265.

Sacks, O. (1990). Seeing voices: a journey into the world of the Deaf. New York, HarperCollins.

Sarimski, K. (2002). Analysis of intentional communication in severely handicapped children with Cornelia-de-Lange syndrome. *Journal of Communication Disorders* 35: 483-500.

Savage-Rumbaugh, E. S., Shanker, S. G. and Taylor, T. J. (1998). Apes, language, and the human mind. New York, Oxford University Press.

Scerif, G., Gomez, J. C. and Byrne, R. W. (2004). What do Diana monkeys know about the focus of attention of a conspecific? *Animal Behaviour* 68: 1239-1247.

Schaller, G. (1963). The mountain gorilla. Chicago, University of Chicago Press.

Schmandt-Besserat, D. (1978). The Earliest Precursor of Writing. *Scientific American* 238(6): 50-59.

Semaw, S. (2000). The world's oldest stone artefacts from Gona, Ethiopia: their implications for understanding stone technology and patterns of human evolution between 2.6-1.5 million years ago. *Journal of Archaeological Science* 27: 1197-1214.

Senghas, A., Kita, S. and Ozyurek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science* 305(5691): 1779-1782.

Seyfarth, R. M. and Cheney, D. L. (1986). Vocal Development in Vervet Monkeys. *Animal Behaviour* 34: 1640-1658.

Seyfarth, R. M. and Cheney, D. L. (1997). Behavioral mechanisms underlying vocal communication in nonhuman primates. *Animal Learning & Behavior* 25(3): 249-267.

Seyfarth, R. M., Cheney, D. L. and Marler, P. (1980). Monkey responses to three different alarm calls: evidence for predator classification and semantic communication. *Science* 210: 801-803.

Seyfarth, R. M., Cheney, D. L. and Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour* 28: 1070-1094.

Shapiro, G. L. and Galdikas, B. M. F. (1999). Early sign performance in a free-ranging, adult orangutan. The mentalities of gorillas and orangutans: comparative perspectives. S. T. Parker, R. W. Mitchell and H. L. Miles. Cambridge, Cambridge University Press: 265-279.

Shepherd, S. V. and Platt, M. L. (2008). Spontaneous social orienting and gaze following in ringtailed lemurs (Lemur catta). *Animal Cognition* 11(1): 13-20.

Sherwood, C. C., Broadfield, D. C., Holloway, R. L., Gannon, P. J. and Patrick, R. H. (2003). Variability of Broca's area homologue in African great apes: Implications for language evolution. *Anatomical Record Part a-Discoveries in Molecular Cellular and Evolutionary Biology* 271A(2): 276-285.

Shettleworth, S. J. (2001). Animal cognition and animal behaviour. *Animal Behaviour* 61: 277-286.

Slocombe, K. E. and Zuberbühler, K. (2005). Agonistic screams in wild chimpanzees (Pan troglodytes schweinfurthii) vary as a function of social role. *Journal of Comparative Psychology* 119(1): 67-77.

Slocombe, K. E. and Zuberbühler, K. (2005). Agonistic screams in wild chimpanzees vary as a function of social role. *Journal of Comparative Psychology* 119(1): 67-77.

Slocombe, K. E. and Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. *Current Biology* 15(19): 1779-1784.

Slocombe, K. E. and Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences of the United States of America* 104: 17228-17233.

Stauffer, R. L., Walker, A., Ryder, O. A., Lyons-Weiler, M. and Blair Hedges, S. (2001). Human and ape molecular clocks and constraints on paleontological hypotheses. *The Journal of Heredity* 92(6): 469-474.

Stoinski, T. S. and Whiten, A. (2003). Social learning by orangutans (Pongo abelii and Pongo pygmaeus) in a simulated food-processing task. *Journal of Comparative Psychology* 117(3): 272-282.

Stout, S. C., Boughner, R. L. and Papini, M. R. (2003). Reexamining the frustration effect in rats: Aftereffects of surprising reinforcement and nonreinforcement. *Learning and Motivation* 34: 437-456.

Tanner, J. E. and Byrne, R. W. (1993). Concealing facial evidence of mood: perspective-taking in a captive gorilla? *Primates* 34(4): 451-457.

Tanner, J. E. and Byrne, R. W. (1996). Representation of action through iconic gesture in a captive lowland gorilla. *Current Anthropology* 37(1): 162-173.

Tebbich, S. and Bshary, R. (2004). Cognitive abilities related to tool use in the woodpecker finch, Cactospiza pallida. *Animal Behaviour* 67: 689-697.

Thieme, H. (1997). Lower Palaeolithic hunting spears from Germany. *Nature* 385(6619): 807-810.

Tomasello, M. (1996). Do apes ape? Social learning in animals: The roots of culture. B. Galef and C. Heyes. San Diego, Academic Press: 319-346.

Tomasello, M. (2003). Constructing a language, HUP.

Tomasello, M. and Call, J. (1997). Primate cognition. New York, Oxford University Press.

Tomasello, M. and Call, J. (2007). Introduction: Intentional communication in nonhuman primates. The gestural communication of apes and monkeys. J. Call and M. Tomasello. Mahwah, NJ, Lawrence Erlbaum Associates, Inc.: 1-15.

Tomasello, M., Call, J. and Gluckman, A. (1997). Comprehension of Novel Communicative Signs by Apes and Human Children. *Child Development* 68(6): 1067-1080.

Tomasello, M., Call, J. and Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour* 55: 1063-1069.

Tomasello, M., Call, J. and Hare, B. (2003). Chimpanzees understand psychological states - the question is which ones and to what extent. *Trends in Cognitive Sciences* 7(4): 153-156.

Tomasello, M., Call, J., Nagell, K., Olguin, R. and Carpenter, M. (1994). The learning and use of gestural signals by young chimpanzees: A trans-generational study. *Primates* 35(2): 137-154.

Tomasello, M., George, B. L., Kruger, A. C., Jeffrey, M., Farrar and Evans, A. (1985). The development of gestural communication in young chimpanzees. *Journal of Human Evolution* 14(2): 175-186.

Tomasello, M., Gust, D. and Frost, G. T. (1989). A longitudinal investigation of gestural communication in young chimpanzees. *Primates* 30(1): 35-50.

Trevarthen, C. and Aitken, K. J. (2001). Infant intersubjectivity: research, theory, and clinical applications. *Journal of Child Psychological Psychiatry* 42(1): 3-48.

van Adrichem, G. G. J., Utami, S. S., Wich, S. A., van Hooff, J. and Sterck, E. H. M. (2006). The development of wild immature Sumatran orangutans (Pongo abelii) at Ketambe. *Primates* 47(4): 300-309.

van Krunkelsven, E., Dupain, J., Van Elsacker, L. and Verheyen, R. F. (1996). Food calling by captive bonobos (*Pan paniscus)*. *International Journal of Primatology* 17(2): 207-217.

van Noordwijk, M. A. and van Schaik, C. P. (2005). Development of ecological competence in Sumatran orangutans. *American Journal of Physical Anthropology* 127(1): 79-94.

van Schaik, C. P. (1999). The socioecology of fission-fusion sociality in orangutans. *Primates* 40(1): 69-86.

van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., Suzuki, A., Utami, S. S. and Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science* 299(5603): 102-105.

Volterra, V., Caselli, M. C., Caprici, O. and Pizzuto, E. (2005). Gesture and the emergence and development of language. Beyond nature-nurture: Essays in honor of Elizabeth Bates. M. Tomasello and D. Slobin. Mahway, New Jersey, Erlbaum.

Wallis, J. (2004). Baboon muzzle wiping: Possibly an ancient gesture of internal conflict. Gestural communication in nonhuman and human primates. Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.

Walraven, V., Vanelsacker, L. and Verheyen, R. (1995). Reactions of a Group of Pygmy Chimpanzees (Pan-Paniscus) to Their Mirror-Images - Evidence of Self-Recognition. *Primates* 36(1): 145-150.

Whiten, A. and Byrne, R. W. (1988). Tactical deception in primates. *Behavioral and Brain Sciences*(11): 233-244.

Whiten, A. and Ham, R. (1992). On the nature and evolution of imitation in the animal kingdom: Reappraisal of a century of research. *Advances in the Study of Behavior* 21: 239-283.

Wich, S. A., Utami-Atmoko, S. S., Setia, T. M., Rijksen, H. D., Schurmann, C. and van Schaik, C. (2004). Life history of wild Sumatran orangutans (*Pongo abelii*). *Journal of Human Evolution* 47(6): 385-398.

Wildman, D. E., Uddin, M., Liu, G. Z., Grossman, L. I. and Goodman, M. (2003). Implications of natural selection in shaping 99.4% nonsynonymous DNA identity between humans and chimpanzees: Enlarging genus Homo.

Proceedings of the National Academy of Sciences of the United States of America 100(12): 7181-7188.

Wimmer, R., Kirsch, S., Rappold, G. A. and Schempp, W. (2002). Direct evidence for the _Homo-Pan_ clade. _Chromosome research_ 10: 55-61.

Yerkes, R. M. and Learned, B. W. (1925). Chimpanzee intelligence and its vocal expressions. Baltimore, Williams & Wilkins.

Zentall, T. R. (1996). An analysis of imitative learning in animals. Social learning in animals: The roots of culture. C. M. Heyes and B. G. Galef. San Diego, CA, Academic Press, Inc.: 221-243.

Zuberbühler, K. (2000). Referential labelling in Diana monkeys. _Animal Behaviour_ 59(5): 917-927.

Zuberbühler, K. (2001). Predator-specific alarm calls in Campbell's monkeys, Cercopithecus campbelli. _Behavioral Ecology and Sociobiology_ 50(5): 414-422.

Zuberbühler, K. (2003). Referential signaling in non-human primates: Cognitive precursors and limitations for the evolution of language. _Advances in the Study of Behavior_ 33: 265-307.

Appendices

Appendix I: IMAGES OF THE ENCLOSURES AT THE THREE ZOOS.

Apenheul outdoor area: two of four islands

Durrell outdoor area: one of two islands

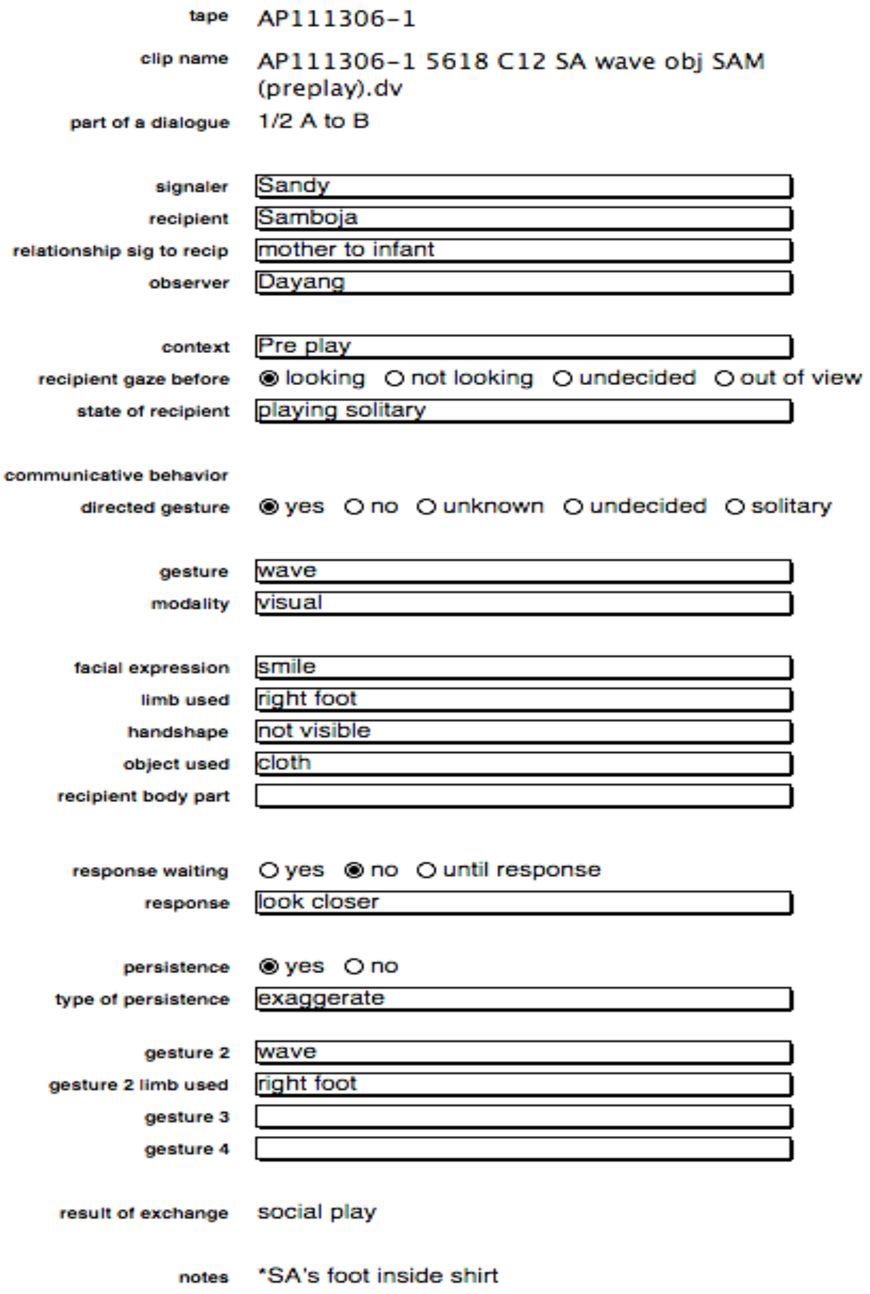
Twycross outdoor area

Apenheul indoor area: one of four rooms (climbing structures extend down to a lower level out of view)

Durrell indoor area: one of two (area extends up approximately 8 meters and is three times the width shown in the photograph)

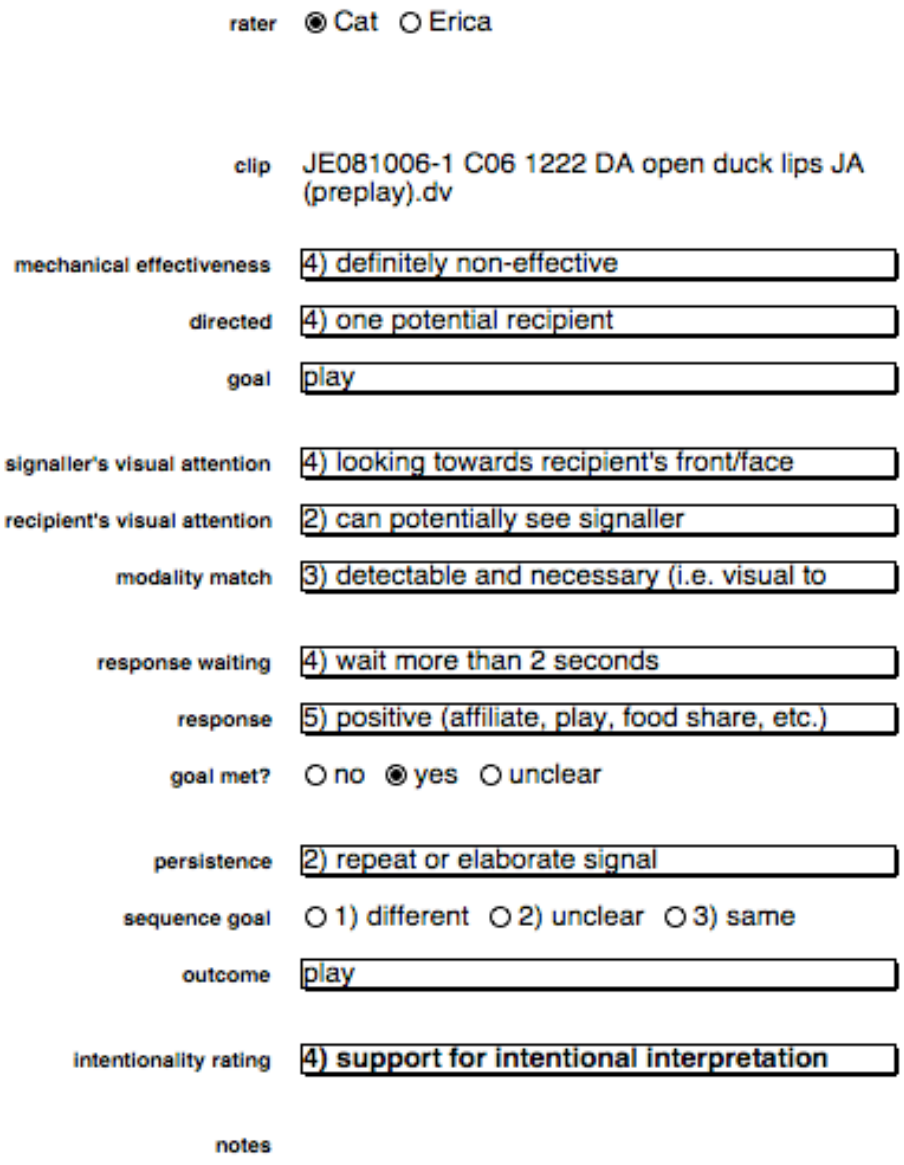
Twycross indoor area: one of two

Appendix II: REPLICATION OF AN ENTRY IN THE FILEMAKER PRO DATABASE USED FOR CODING GESTURES



tape	AP111306-1
clip name	AP111306-1 5618 C12 SA wave obj SAM (preplay).dv
part of a dialogue	1/2 A to B
signaler	Sandy
recipient	Samboja
relationship sig to recip	mother to infant
observer	Dayang
context	Pre play
recipient gaze before	● looking ○ not looking ○ undecided ○ out of view
state of recipient	playing solitary
communicative behavior	
directed gesture	● yes ○ no ○ unknown ○ undecided ○ solitary
gesture	wave
modality	visual
facial expression	smile
limb used	right foot
handshape	not visible
object used	cloth
recipient body part	
response waiting	○ yes ● no ○ until response
response	look closer
persistence	● yes ○ no
type of persistence	exaggerate
gesture 2	wave
gesture 2 limb used	right foot
gesture 3	
gesture 4	
result of exchange	social play
notes	*SA's foot inside shirt

Appendix III: REPLICATION OF AN ENTRY IN THE FILEMAKER PRO DATABASE USED FOR INTER-OBSERVER RELIABILITY CODING



rater	● Cat ○ Erica
clip	JE081006-1 C06 1222 DA open duck lips JA (preplay).dv
mechanical effectiveness	4) definitely non-effective
directed	4) one potential recipient
goal	play
signaller's visual attention	4) looking towards recipient's front/face
recipient's visual attention	2) can potentially see signaller
modality match	3) detectable and necessary (i.e. visual to
response waiting	4) wait more than 2 seconds
response	5) positive (affiliate, play, food share, etc.)
goal met?	○ no ● yes ○ unclear
persistence	2) repeat or elaborate signal
sequence goal	○ 1) different ○ 2) unclear ○ 3) same
outcome	play
intentionality rating	**4) support for intentional interpretation**
notes	

Appendix IV: INTENTIONAL GESTURES OF EACH INDIVIDUAL.

	Apenheul											Durrell								Twycross								
	Binti	Dayang	Fin	Jos	Karl	Katja	Radja	Samboja	Sandy	Silvia	Willie	Dagu	Gempa	Gina	Jaya	Jiwa	Julitta	Mawar	Putri	Batu	Kibriah	Maliku	Miri	Satu	Tamu	Theodora	Tiga	**Total**
AGE	J	I	A	A	A	J	A	I	A	A	J	A	I	A	J	J	A	A	I	A	A	A	I	J	I	A	J	
SEX	F	F	F	F	M	F	F	F	F	F	M	M	M	F	M	M	F	F	F	M	F	F	F	F	F	F	M	
air bite	1					2		1			1				1									2	1		5	14
air grab	2		2			2			2		3				1							1				1		14
arms up	1										4												5				7	17
back roll											1							2									5	8
beckon														1									1				1	3
bite		2	1	1		1		6	1		10			1	4			2			1	4	2				2	38
brush	1			2			1	1	4		1		1	2							3	2		2		2	3	25
cover											1		2														5	8
dangle		1	2			7		2			10		2		1	1		1	1								1	29
drag object															4													4
duck lips closed									1			3		1														5
duck lips open												11																11
embrace				2		1			2		1			1	1						2		3		1		7	21
embrace pull									4								1	1			1							7
fake																										9		9
food beg orally	2							2			1				4								2				2	13
frog lips									1																		1	2
grab	3	2		1		2		3	2		12		4	1	4		1	6			1	5	2		2	3	14	68
grasp		5				4		3	3	1	8		8		3	2	1	7			3		2	2	2	5	7	66
headstand																1		2								2	4	9
hit		2	2	1		1		12	2		13			3	1	1	1	4			6		5			1	13	68
hit ground/obj	1	2				4		5	5		4	3					1					4	1	2		2	1	35
hold hand				1					4																1		1	7
kiss									1		1			2														4
lip smile (tense)	8					1																						9
look back														3				2			1							6
mouth												2			5												4	11
nudge			2	8	2				1	1				6		1	3	1			4			2		2	1	34
offer				1					1					3	2		1										1	9
peer			2	2		3		1			4		2		9				1				2				2	28
play face													1					1				1				2		5
poke											1		1										2					4

	Apenheul											Durrell								Twycross								
	Binti	Dayang	Fin	Jos	Karl	Katja	Radja	Samboja	Sandy	Silvia	Willie	Dagu	Gempa	Gina	Jaya	Jiwa	Julitta	Mawar	Putri	Batu	Kibriah	Maliku	Miri	Satu	Tamu	Theodora	Tiga	**Total**
AGE	J	I	A	A	A	J	A	I	A	A	J	A	I	A	J	J	A	A	I	A	A	A	I	J	I	A	J	
SEX	F	F	F	F	M	F	F	F	F	F	M	M	M	F	M	M	F	F	F	M	F	F	F	F	F	F	M	
pout																					1						4	5
present body part												1												1				2
present genitals				6		1		1																			1	9
pull	2	2		1	2	4		9	6		11		1	1	5		4	7			4	6	2		1		44	112
pull away	0		4						3																	1		8
pull away appen.	1			1		1			2					1	1										1	2		10
pull hair		5	2			4		3			7				3	2						2	2		5		1	36
push	2		1	3			2	3	3		1	1		1	4			7	1		7	1	1	2	1	3	14	58
put obj. on head									2		1	2															2	7
raise arm	1				1	2			1	2					1	1					4	1	4					18
rasp-berry face						3			1																	1		5
reach				1		1		7	5		2	3	4	1	2		1	2			1	1	6	3	6		3	49
restrain									3																			3
roll on back											2		1		3	1							2	1			4	14
seize								4	3		1		2	2	11			1					3		2	2	12	43
shake		2						1			3				1							1	1					9
shoo			2		2				2					6		1		1			3					2		19
show				3											1													4
shrug																										7		7
simultaneous hit				1		1															2	4					1	9
somersault													1			1							7		1		32	42
swat			1	1				1	1		9		3		6						2	4	2			3	4	37
swing			1					1			2				4								2	1				11
tandem walk																						1					17	18
tap				2		3					1		2	1	1		1					1		1		5	2	20
teeth bared						1															1							2
tense upper lip									1												1							2
tongue out																					3					1		4
touch	4	4	1	9	2	7		10	4		12	3	5	2	13		3	5	2	1	2	2	5	3	2	6	13	120
turn head	2					2		2	3		2			1	10									1			2	25
wave			1	3		4		1	3		7		4		4								4	1			1	33
whistle face										1				1														2
Total	**31**	**27**	**24**	**50**	**9**	**62**	**3**	**79**	**77**	**5**	**137**	**29**	**44**	**41**	**110**	**12**	**18**	**52**	**5**	**1**	**53**	**41**	**68**	**24**	**26**	**62**	**244**	**1334**

Appendix V: SUMMARY TABLE OF NUMBERS OF GESTURES IDENTIFIED, DATASETS USED, AND ANALYSES PERFORMED ON THE DIFFERENT DATASETS.

SUMMARY	71 *potential* gesture types (44 visual, 24 tactile, 3 auditory) - 9 potential gesture types excluded for lack of intentional use - 1 auditory gesture type (hit ground/object) reclassified as visual - 2 new gestures (seize and pull away appendage) identified 64 *intentional* gesture types (39 visual, 25 tactile)

Data set	Response	No response	Analyses
1581 potential gesture tokens	985 (62%)	596 (38%)	- Intentionality criteria applied to this set to identify communicative gestures (Chapter 3.2)
1334 intentional gesture tokens (earning an intentionality rating of 3 or 4)	832 (62%)	502 (38%)	- Gesture use by individuals and groups (Chapter 3.6) - Effect of age of signaller (Chapter 3.7)
1421 broad definition intentional gesture tokens (intentional solitary gestures plus all gestures in sequences containing intentional gestures)	876 (62%)	545 (38%)	- Sensitivity to gaze (Chapter 3.3)
698 gestures with observable goals	536 (77%)	162 (23%)	- Gesture meaning (Chapter 3.4, 3.5)
730 gestural utterances			- Utterance length (Chapter 4.3)
1042 gestures in sequences (in 349 sequences)	570 (55%)	472 (45%)	- Sequence meaning (Chapter 4.3) - Persistence rates based on recipient response (Chapter 4.4) - Types of persistence (Chapter 4.5) - Sensitivity to what recipient has witnessed (Chapter 4.6)
381 solitary gestures	312 (82%)	69 (18%)	