DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„Approaching cross-modal perception in kea (*Nestor notabilis*)“

verfasst von / submitted by

Michael Klosz

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magister der Naturwissenschaften (Mag.rer.nat.)

Wien, 2017 / Vienna, 2017

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 190 313 445
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Lehramtsstudium UF Geschichte, Sozialkunde und Politische Bildung UF Biologie und Umweltkunde
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Ludwig Huber

Sworn Declaration

I hereby declare that I wrote this diploma thesis on my own and that I did not use any other than the given sources. All content drawn from external sources are indicated as such. The thesis has not been submitted to any other examining body and has not been published.

Acknowledgements

I want to thank Ludwig Huber for giving me the possibility to do a study in this fascinating field of research and for supervising my diploma thesis, as well as Raoul Schwing for introducing me to the amazing kea, for statistical support, for helpful comments during the whole project and, last but not least, for insightful conversations. Also, I want to thank Amelia Wein for sharing her experience in work with this species, and Mark O'Hara for helping me with the touchscreen program. Moreover, I want to thank the whole team at the Haidlhof research station for their great work. Special thanks go to Marlene, and also to friends & family for the support during the work for this thesis.

Table of contents

Approaching cross-modal perception in kea (*Nestor notabilis*)

1. Abstract

Cross-modal perception is the capacity to abstract and exchange information between different sensory modalities. It is beneficial in many situations and has been shown to exist in a variety of animal species. The kea (*Nestor notabilis*), a parrot from New Zealand, is known for its play behaviour, its technical intelligence, its ability to learn and its sociality. All of these features allow indications on the ability to exchange information across modalities and to integrate it. After a literature review on cross-modal perception among animals, two different experiments will be presented, which were conducted to investigate cross-modal abilities in kea. The first was an audio-visual matching-to-sample task on a touchscreen, showing pictures of familiar objects that should be matched to an audio sample of a recording of the sounds produced by one of the objects. As the birds were unable to produce significant results a second task was designed, to overcome possible obstacles of the first experiment. Here the test subjects were confronted with a two-choice discrimination task, but now with real objects. Two differently shaped wooden boxes were presented, both of which with a specific sound, coming out of one of the boxes. Depending on which sound was playing the correct box should be chosen. All of the tested kea finished the training for the second experiment successfully, but testing then had to be stopped due to time constraints. So far we could show that kea can use auditory cues in a multimodal experiment. Building on these results we will recommend steps for future research on cross-modal perception in kea.

Deutsche Zusammenfassung

Kreuzmodale Wahrnehmung ist die Fähigkeit Informationen zwischen verschiedenen Sinnesmodalitäten auszuschauen und zu abstrahieren. Sie scheint in vielen Situationen vorteilhaft zu sein und wurde in einer Vielzahl an Spezies nachgewiesen. Der Kea (*Nestor notabilis*), ein Papagei aus Neuseeland, ist bekannt für sein Spielverhalten, seine technische Intelligenz, seine Lernfähigkeit und seine Sozialität. All diese Eigenschaften weisen auf die Fähigkeit hin Informationen zwischen Modalitäten auszutauschen und sie zu integrieren zu können. Nach einem Forschungsüberblick über kreuzmodale Wahrnehmung unter Tieren werden zwei verschieden Experimente erklärt werden, die durchgeführt wurden um kreuzmodale Fähigkeiten des Kea zu untersuchen. Das erste war eine audio-visuelle ‚matching-to-sample' Aufgabe an einem Touchscreen, auf dem Bilder von bekannten Objekten gezeigt wurden, die mit einem Audiosample eines aufgenommenen Geräusches, das von einem der Objekte erzeugt wurde, übereingestimmt werden mussten. Da die Vögel unfähig waren signifikante Resultate zu produzieren wurde eine zweite Aufgabe konzipiert, um mögliche Schwierigkeiten des ersten Experiments auszuschließen. Hier waren die Testsubjekte mit einer ‚two-choice-discrimination' Aufgabe konfrontiert, nun aber mit echten Objekten. Zwei unterschiedlich geformte Holzboxen wurden präsentiert, beide mit einem bestimmten Ton, der jeweils aus einer der Boxen kam. Je nachdem welcher Ton gespielt wurde sollte die richtige Box ausgewählt werden. Alle getesteten Kea beendeten das Training für das zweite Experiment erfolgreich, danach musste die Untersuchung aber wegen Zeitbeschränkungen beendet werden. Bisher konnten wir zeigen, dass Kea Audiohinweise in einem multimodalen Experiment verwenden können. Darauf aufbauend werden wir Schritte für zukünftige Forschung an kreuzmodaler Wahrnehmung von Kea vorschlagen.

2. Introduction

2.1. General introduction to cross-modal perception

The world is a conglomeration of information, transported over diverse media in a variety of encodings. Through our sensory systems, we are capable of retrieving parts of this continuous bombardment of signals, filtering useful information out of all the noise. We can see, hear, smell, taste and feel certain properties of our environment, and only these capabilities enable us to react properly to events and to act purposefully in the world. Huber (2000) notes that it is through linking of these perceptual concepts from different modalities that global, meaningful perception emerges.

Most objects and events "provide multimodal stimulation and evoke a diversity of visual, auditory, tactile, and olfactory impressions simultaneously" (Lickliter and Bahrick 2000). These are processed through distinct perceptual systems, but here the question arises how "objects and events [are] experienced as unitary when they stimulate receptors that give rise to different forms of information", in other words "how [] different modes of sensory stimulation [are] bound together" (Bahrick et al. 2004). This 'binding problem' is hotly debated, as it is still not thoroughly understood how different qualities of the world, which are separated as they enter the 'doors of perception', are transformed and connected into a behaviourally usable, coherent percept.

The illustration from Meredith (2002) (see Fig. 1) nicely depicts this process of detecting and integrating distinct physical information properties, which are being produced or reflected in the environment. These properties include light (for vision), airborne oscillations (for audition), solid-body vibrations (for sensation) and distribution of air- or waterborne molecules (for olfaction and gustation), all of which do not affect each other directly, and are accordingly detected as different stimuli by distinct sensory systems with specific receptors. Multimodal events or objects, which emit more than one of these information properties, can only be perceived as such through a process of integration across modalities within a sensory and cognitive system, thus through

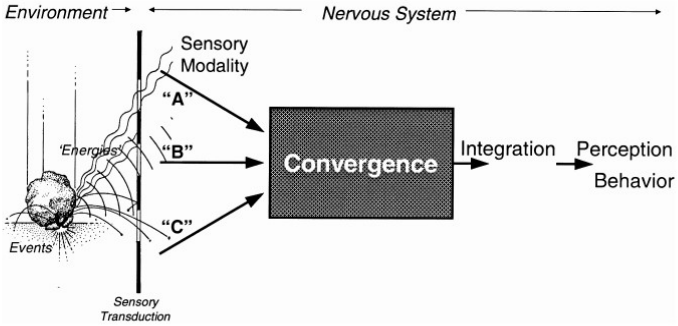

Figure 1: Multimodal convergence as depicted by Meredith (2002): separated information properties from the environment get integrated to produce a coherent percept

cross-modal perception (henceforth abbreviated as CMP). CMP has consequently been described as "the capacity to abstract and exchange information between different sensory modalities" (Davenport et al. 1973), which will also be our working definition. In the literature this capability has also been called intermodal generalization (Davenport 1977), the interrelatedness of the senses (Ibid.), binding (Roskies 1999) and intersensory perception (Lickliter and Bahrick 2000). The perception of multimodal events is therefore dependent as much on the properties of the system that is detecting and integrating it, as it is on the perceived event itself.

Only this CMP enables us humans to experience the richness of the world in the way that we do. It gives us, for example, the combined sensations of seeing, smelling and tasting good food, or of enjoying the smell of salty air near the sea, while hearing waves break and feeling a soft breeze on our skin. As biologists, however, we should not only appreciate these human experiences, but also ask if animals have similar sensations and how they experience the world. CMP has long been seen as an "exclusive property of man" (Davenport et al. 1973), but has since been shown to exist in a variety of animal species. A possible reason for this former misconception might have been that for some time it was proposed that language is a necessary condition for CMP, leading to the conclusion that animals do not perceive cross-modally, as they have no language (Davenport 1977). This notion, however, is outdated. It seems much more plausible that it is the other way around, meaning that CMP is a necessary (but not sufficient) prerequisite for language acquisition (Ettlinger and Wilson 1990; Hashiya and Kojima 2001). Another possible reason for denying animals the capability for CMP is that "this capacity lies at the heart of the questions posed by physiologists and psychologists concerning the unity or integration of the senses" (Davenport et al. 1973), and that is somewhat related to the controversial problem of consciousness. Indeed, Roskies (1999) links the processes for binding information across modalities to the "mystery of consciousness", considering if similar mechanisms are responsible for "the unity of phenomenal experience". Understandably, she concludes that the binding problem is "one of the most puzzling and fascinating issues" in the cognitive sciences.

Claims about human uniqueness have often been made before relevant studies investigated the presence of certain concepts among other animals. Such studies include investigations of culture or tool use, both of which have by now been shown to exist (to some extent) in quite a few animal species (Boesch 2007). The same applies for CMP. Research results from several species suggest that they indeed are capable of perceiving cross-modally (see next section). According to Davenport (1976) there are two main approaches to investigate CMP in animals: cross-modal transfer and cross-modal matching-to-sample. In cross-modal transfer tasks, the subjects are trained to discriminate stimuli in a single sensory modality. The same discrimination is then tested in another modality. A transfer happened when the second discrimination is learnt faster than the first, if not immediately. An example would be for subjects to learn to visually discriminate a sphere from a cube, and then to test if this can also be done haptically (Ettlinger and Wilson 1990).

In cross-modal matching-to-sample tasks the subjects need to match a sample in a given modality with the equivalent one of several stimuli in another modality, such as matching an object, which can be touched but not seen, with the one photographs of several that is depicting the S+. The presentation of sample and stimuli can be simultaneous or delayed (Davenport et al. 1975). The ability to match sample and stimulus is often also measured indirectly through looking preference (e.g. Sliwa et al. 2011) or violation of expectation designs (e.g. Adachi et al. 2007).

Ettlinger and Wilson (1990) additionally distinguish cross-modal recognition tasks as a separate approach to test CMP, but the used methods are just a variation of matching-to-sample tasks. The distinguishing feature of recognition tasks is that the associations that are being tested are not newly learnt, but already existent, for example when identifying familiar individuals based on the sounds they produce (e.g. Kondo et al. 2012).

2.1.1. Research history

The philosophical discussion of topics related to CMP goes back to the 17th/ 18th century, at least, conducted for example by Locke and Berkeley (Davenport 1976). The experimental history, however, started only in the 20th century. The earliest empirical study dates back to 1932, investigating visual-olfactory transfer in a fish species, the European minnow, *Phoxinus laevis*, (Schiller 1932). The term 'cross-modal' was introduced in 1960 by Ettlinger, for a study on cross-modal transfer in rhesus monkeys, *Macaca mulatta* (Ettlinger 1960; Ettlinger and Wilson 1990). In the following years, several papers reported on visual-haptic CMP in primates (e.g. Ettlinger and Blakemore 1967; Davenport et al. 1973; Ettlinger 1973; Cowey and Weiskrantz 1975; Malone et al. 1980). Research methods, as well as modalities examined, were extended in subsequent years. CMP, which had before been described as an “exclusive property of man” (Davenport et al. 1973) or at least as restricted to humans and apes (Ettlinger 1973), has since been investigated and shown to exist in many other species. Besides primates, examples are rats, *Rattus norvegicus* (Over and Mackintosh 1969), rabbits, *Oryctolagus cuniculus* (Yehle and Ward 1970), dogs, *Canis lupus familiaris* (Adachi et al. 2007; Ratcliffe et al. 2014), horses, *Equus caballus* (Proops et al. 2009; Lampe and Andre 2012a) and also large-billed crows, *Corvus macrorhynchos* (Kondo et al. 2012). Especially in recent years the focus shifted from CMP of objects to that of con- and heterospecific individuals (e.g. Adachi and Fujita 2007; Adachi et al. 2009; Proops et al. 2009; Sliwa et al. 2011; Kondo et al. 2012; Lampe and Andre 2012b; Proops and McComb 2012).

Abilities incorporating several sensory modalities have not only been tested using behavioural tasks, but also through other means. Research on multisensory perception in primates, for example, has been extended to the neuronal level, where specific regions, with neurons that respond to multisensory stimuli have been found (Romanski 2007). In zebra finches, *Taeniopygia guttata castanotis*, it has been shown that visual stimuli can activate auditory brain regions (Bischof and Engelage 1985). Studies on cats, *Felis catus*, hint at neurons that respond only to multi-modal, but not to unimodal stimulation (Partan and Marler 1999). Furthermore Lickliter and Bahrick (2000, 2004) propose that certain neurons show more activity during multimodal stimulation than one would expect from adding the neural consequences of stimuli from each modality alone. This “’multiplicative’ effect of multimodal stimulation” on the neuronal level leads to enhanced behavioural responsiveness, which is thought to increase attentiveness, and might also enhance the ability to detect and recognize objects. Coordinated multimodal stimulation also seems important for development, even in the prenatal period (King et al. 2001; Sakata et al. 2004)

Learning experiments have contributed to the understanding of CMP as well: it has been shown in chickens, *Gallus gallus domesticus*, that the simultaneous presentation of visual and auditory stimuli enhances the learning effect in both modalities (Van Kampen and Bolhuis 1991, 1993), and that they prefer imprinting objects that involve stimuli from more than just the visual modality (Smith and Bird 1963; Van Kampen and Bolhuis 1992). In nightingales, *Luscinia megarhynchos*, synchronous visual stimulation in addition to the presentation of sound recordings enhanced song learning (Hultsch et al. 1999). Rowe (2002) showed a similar effect the other way around. In his experiment, an additional auditory stimulus enhanced the performance in visual discrimination learning in chickens.

All these studies present examples of cross-modal interaction in a variety of animals, where the neuronal, behavioural or learning response of one modality is influenced by the stimulation of another modality.

2.1.2. Adaptive value of CMP

If CMP is as wide spread in the animal kingdom as it should be based on theory and recent research, then the question arises what its biological advantage is. Factors of utter importance for evolutionary change are the social and ecological environment of a species (Lefebvre et al. 2004). Every animal needs to respond appropriately to its mates, rivals, juveniles and predators, but also to its food and potential dangers in its surroundings, to be successful in the 'struggle for life'. Here CMP might in several respects be beneficial compared to unimodal perception. An obvious advantage is that it increases the salience of stimuli, increases reliability of perception and reduces perceptual ambiguities (Lalanne and Lorenceau 2004; Spence 2011). For example, vision and audition often lead to different but complementary information, as is the case in stimulus localization. While vision has a high spatial resolution, but is restricted to the visual field and only works in the presence of light, audition helps to locate stimuli in any direction, also behind obstacles, and also works in the dark. Combining the benefits of several modalities therefore enhances accuracy and reliability in perception (Knudsen and Brainard 1995). Synchronizing of this complementary information is also important for the development of sensory systems. Cross-modal integration is, for example, thought to be necessary for calibrating the auditory landscape through vision (Knudsen and Brainard 1995).

Another adaptive value of CMP can be found in social behaviour. Especially in individualized societies specialized information processing has been suggested, to allow recognition, long-term relationships and tracking others' social relationships (Emery 2006). Cross-modal abilities could be an example of these types of processing. They are, however, not only useful for recognition of individual conspecifics, but also of kin in general, of gender, of other species, of predators and of prey. It is also not only beneficial to recognize others, but to be recognized. Advantages of identity signalling are increased altruism from conspecifics, stability in reciprocal interactions, decreased risk of inbreeding and less aggressive competition over status (Tibbetts and Dale 2007).

Also, communication might profit from CMP, especially in group-living, social species (Partan and Marler 1999). Many biological displays are complex and include signals from more than one modality. Such signals are used in mating, begging and parent-offspring communication, but also as warning displays and in aggressive situations. Investment in this additional effort (compared to unimodal signals) might pay off as better detectability, discriminability and memorability of signals (Rowe 1999). These advantages have been suggested to be "powerful selective forces in signal design" and signallers will be selected to produce such easily receivable signals (Guilford and Dawkins 1991).

Signals incorporating more than one modality might further function as "backup signals", ensuring that the signal is received, even when facing environmental noise, or might serve as "multiple messages", signalling different qualities. Also, both explanations might play a role, as they are hard to isolate from each other (Johnstone 1996). Explaining this idea in more detail, Partan and Marler (1999) (see Fig. 2) sum up different possibilities for the outcome of multimodal integration. Redundancy means that through different channels the same information is transported, whereas nonredundancy means that the signals transport different information. The "intersensory redundancy" hypothesis of Lickliter and Bahrick (2000) also addresses this topic and claims that if the same information is presented through different sensory modalities, it elicits an increased response compared to unimodal signals ("Enhancement" in the illustration). In reviewing (human) psychological studies, Ernst and Bülthoff (2004) come to the conclusion that the value of multimodal perception is to maximize information (through combination of nonredundant signals) and to increase reliability of perception (through integration of redundant signals).

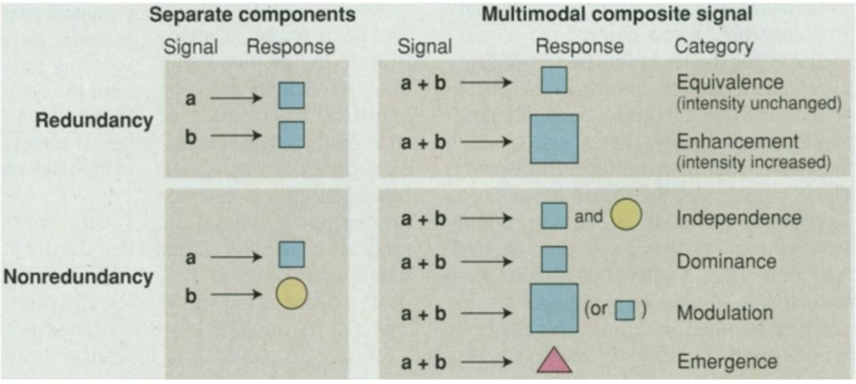

Figure 2: Different possible outcomes of multimodal integration, as described by Partan and Marler (1999): depending on whether the signals from different modalities are redundant (same information is transported) or nonredundant (different information is transported) the responses can vary

So far, CMP has been discussed with respect to conspecifics, but also recognition of predators and potentially dangerous prey can be improved by it. Some birds, for example, are confronted with aposematic insects that not only use warning colours, but also produce sounds or odours when being attacked. Through integrating information from these different modalities (possibly also gustatory perception, indicating toxicity), detectability and recognition of inedible or potentially dangerous insects is increased, which facilitates avoidance of them (Rowe and Guilford 1999; Rowe 2002). The same applies for being alert from predators. Using auditory or olfactory cues before a predator appears visually can be of vital importance. But again, not only receiving signals, but also sending them can be important here, as when using multimodal signals to deter potentially dangerous animals (Rowe 1999).

2.1.3. Neuroscience of CMP

Different information properties of multimodal events are distinct and also the sensory systems detecting them are separated themselves. Within a neural system these properties then need to be connected somehow, to allow CMP, which we have defined as "the capacity to abstract and exchange information between different sensory modalities" (Davenport et al. 1973). For this process of integration across modalities two general approaches can be found in the literature, illustrated by the depiction from Meredith (2002) (see Fig. 3). The model of areal convergence assumes that single neurons only respond to a certain modality and do not integrate input from different sensory systems. Integration only occurs in a given area, which could be achieved through temporal synchronisation of neuronal activity, through which different perceptual features might be connected, resulting in a unified percept (Huber 2000). The neuronal convergence model, on the other hand, assumes that the same neuron integrates information from more than one modality, and that it responds to stimulation from several sensory systems (Lickliter and Bahrick 2000, 2004). As illustrated in

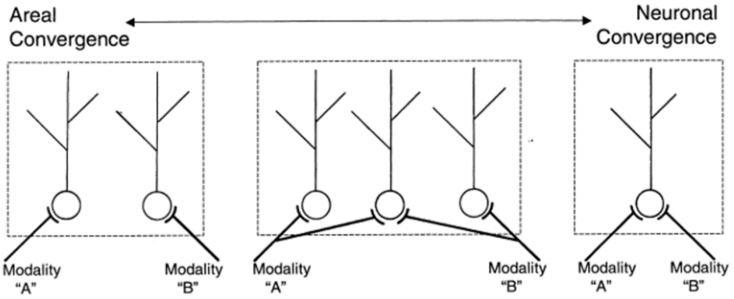

Figure 3: Models for multimodal convergence by Meredith (2002): in the areal convergence model neurons only respond to input from a single modality, integration occurs in a given area; in the neuronal convergence model single neurons integrate input from more than one modality; in the hybrid model (middle) both forms of integration occur

the central picture also a continuum between the two models, incorporating features of both, is possible, which Meredith (2002) calls the most likely approach.

The question now arises where these areas of convergence might be located. In our considerations we will focus on avian brains, especially those of parrots, which have (together with those of corvids) been compared to primate brains, due to their large relative brain and telencephalic volumes (Iwaniuk et al. 2005; Schloegl et al. 2009). But not only have the brains as a whole been related to each other, also specific areas have been compared. So has, for example, an avian equivalent to the primate neocortex been described: the avian pallium, which is involved in sensory processing, learning, memory and attention (Shimizu 2009). Emery (2006) even suggests a functionally equivalent area to the prefrontal cortex, namely the caudolateral nidopallium (CDNL, also NCL; the nidopallium was formerly known as neostriatum), which is associated with solving complex cognitive tasks. Additionally, this area might play a role in CMP, as it is connected reciprocally with secondary sensory regions of all modalities. In the same way Shimizu (2009) points to the fact that cells in the nidopallium receive afferent projections from visual, auditory and somatosensory areas. The structure containing the nidopallium is larger in parrots and corvids, compared to other birds. Summing up, the addressed area has been called “a multimodal forebrain area that is located at the transformation from sensation to action”, distinctive for the bird brain (Güntürkün 2005).

The comparison of parrots (and corvids) with apes is an example of “convergence in cognitive behaviour despite divergence in the neuronal substrate” (Watanabe and Huber 2006), reflecting similar evolutionary adaptations, which might be the result of the “need to solve comparable social and ecological problems”. When taking a broader perspective, one can even specify six socio-ecological factors that are correlated with complex cognition in a variety of animals (corvids, parrots, monkeys, apes, elephants and cetaceans): “omnivorous generalist diet, highly social, large relative brain size, innovative, long developmental period, extended longevity and variable habitat” (Emery 2006), all of which is the case for the kea, the species investigated here.

2.2. The kea as a promising test subject to investigate CMP

The kea (*Nestor notabilis*) is a parrot endemic to New Zealand, with characteristic olive green plumage, under which orange-red and blue feathers are hidden that get revealed during flight or social displays. It is about 46-48 cm tall and weighs between 600 g (females) and 1000 g (males). It belongs to the genus Nestor in the family Strigopidae, in the superfamily of the Strigopoidea, in the order of the Psittaciformes (parrots). Kea (singular and plural are the same, as it derives from Mâori, which does not distinguish between these forms (Schwing 2014)) live and breed in the alpine environment of the Southern Alps of the South Island, New Zealand (Del Hoyo et al. 1997).

The kea is known for is its technical intelligence and its curious, playful, destructive and therefore (from a human point of view) often mischievous behaviour. This behaviour is a reflection of the kea's cognitive adaptability and enables it, in its harsh environment with fluctuating resources, to explore its surroundings thoroughly and makes it an opportunistic forager (Watanabe and Huber 2006). With its varied diet the kea is beyond doubt a textbook example of an omnivorous generalist. It not only digs for roots, catches insects and eats eggs, but has also learnt to make use of human refuse and to feed on sheep (which nearly led the kea to be hunted to extinction) (Diamond and Bond 1999). Apart from this ability to adapt to available resources, it is remarkable to what extent this bird is able to recognize and remember usable food sources. It forages more than 100 species of plants and even knows which parts of them to be edible (Huber and Gajdon 2006).

Many behavioural and choice tasks have investigated the cognitive abilities of these birds. They have been shown to have the abilities for emulation learning (Huber et al. 2001; Huber 2007), means-end comprehension (Werdenich and Huber 2006; Auersperg et al. 2009), tool use in experimental situations (Auersperg et al. 2011), second-order tool use (Auersperg et al. 2010), discrimination between and generalization of high-level patterns (Stobbe et al. 2012), reversal learning (O'Hara et al. 2015), and inference by exclusion (O'Hara et al. 2016). Huber and Gajdon (2006) conclude that kea might challenge primates not only in telencephalic volume, but also in cognitive abilities and in some tasks have even been shown to do better than chimpanzees (Auersperg et al. 2009).

Only few studies have investigated perceptual features of this species, and if so only in unimodal tasks (vision: Wein et al. 2015; O'Hara et al. 2015; audition: Schwing 2014; Schwing et al. 2016; olfaction: Gsell et al. 2012). We have no scientific knowledge of cross-modal perceptual abilities of the kea, even though several characteristics of the kea suggests their existence. The kea's sociality and the hierarchies in kea populations indicate a high level of individual recognition of conspecifics, as has been proposed for a range of species (Tibbetts and Dale 2007). It seems reasonable that this ability is not only limited to the visual modality. Similarly communication between kea would profit from CMP, as is indicated by several displays, which involve visual (posture, head feathers, wings) and auditory (calls) signals, for example warning displays. Also, the experience of working with kea on a daily basis suggests that they are able to recognize humans cross-modally (from visual and auditory, maybe even from olfactory cues). This manifests itself in individual relationships with and reactions to people, which is also the essence of many anecdotes. Regarding the abovementioned large repertoire of food plants, it has been suggested that the kea might possess "a large toolkit of perceptual and olfactory skills" to remember them (O'Hara et al. 2012) and that the kea's "diet is the result of essentially trying everything and keeping what works" (Huber 2007). Such an efficient recognition system would profit from incorporating not only visual, but also haptic, olfactory, gustatory and maybe even auditory information, thus from CMP. Finally, one of the defining characteristics of the kea, its playfulness, could also be connected with CMP. Through continuous interactions with an object, a lot of information is provided though several modalities and it seems plausible that the 'player' acquires (foremost visual and haptic, but also acoustic) knowledge about these object features. The kea, due to its explorative and manipulative behaviour, seems to be a promising species to investigate CMP in.

Despite a lot of research on CMP over the last decades, this comment from Davenport is as accurate as it was in 1976: "Unfortunately, we know very little about the phylogenesis of cross-modal perception, since so few taxa have been studied in so few situations" (Davenport 1976). To shed light on the evolutionary history of CMP, the kea is a very interesting species to investigate, due to its phylogenetic position. It belongs (together with the kaka, *Nestor meridionalis*) to the most basal parrots and is phylogenetically closest to the parrots (Psittaciformes) last common ancestor with the falcons (Falconidae) (Jarvis et al. 2005; Schwing 2014). With this thesis, we hope to initiate more studies on CMP in kea and to make a contribution so that Davenport's comment can be retired, finally.

2.3. Research question & hypothesis

All characteristics of this species mentioned so far raise the question if the kea is able to build cross-modal representations of objects and of con- and heterospecifics. A first step towards answering these and potential follow-up questions is to investigate cross-modal associations in kea in general. Kea have successfully solved several visual tasks, using a touchscreen (e.g. O'Hara et al. 2015, 2016; Wein et al. 2015). By integrating auditory information in our experimental design, we made the task more complex and thus more difficult to solve, but it has been shown that kea are able to use sounds to solve a spatial discrimination task (Schwing et al. 2016). Therefore, we conclude that the ability to discriminate objects on pictures (Wein et al. 2015) can be extended to discriminate pictures of objects based on different sounds. The first study consisted of an audio-visual matching-to-sample task on the touchscreen, showing pictures of familiar objects, while playing a sound that is produced by one of the objects.

After this study design provided no significant data (as will be described below) and we could not be sure whether this showed a perceptual or cognitive deficit, or was due to the experimental setup, we decided to simplify the design by using real objects instead of pictures, as O'Hara et al. (2015) showed that discrimination tasks in kea work better with real objects. In this second design, we used two wooden boxes, shaped as a cube and a pyramid, both of which were presented with a certain sound.

The research question for both experiments was:

- ***Can kea learn to associate a visual and an auditory stimulus, and then choose the correct stimulus in an audio-visual matching-to-sample task?***

Our hypothesis was:

- ***Kea do have the capability for CMP and therefore are able to solve the audio-visual matching-to-sample task.***

Due to time constraints, we were (after abandoning the first design) only able to finish the training for the second experimental design. Of the 11 birds we started with two were excluded from the sample because they stopped participating. The remaining nine kea finished training successfully. Building on these results, we will recommend steps for future research on CMP in kea.

3. General material & methods

3.1. General Procedure

All experiments described here were conducted from May 2016 until February 2017. There were between one and five days of testing per week. Each individual was given one to two sessions per day (if tested twice: once in the morning, once in the afternoon) and each session consisted of 20 trials. The training criterion for both designs was set to 18 correct out of the 20 trials (90%) over two consecutive sessions, after which the subject would be transferred to the next phase. Successful trials were rewarded with an eight of a peanut. In the testing compartment, the subjects were visually isolated from the other birds by an opaque sliding wall.

The experiments were strictly non-invasive and were stopped if a subject refused to participate for 20 minutes. Also, no bird was forced in any way to take part and could walk away from the testing apparatus any time if not interested. Upon repeated refusal to participate (three consecutive sessions), the individual was dropped from the list of test subjects.

3.2. Housing

The kea in this study were housed in an outdoor aviary (52m * 10m * 6m) at the Haidlhof Research Station (cooperation between the Department of Cognitive Biology at the University of Vienna and the Messerli Research Institute at the University of Veterinary Medicine Vienna) in Bad Vöslau. The aviary was built according to standards acquired over a decade of studies on the cognition and behaviour of the species, containing breeding cabins, foraging tables, tree trunks, perches, wooden shelters, two water ponds, flower beds, rocks, and other objects for enrichment. It could be divided into 9 compartments. The floor of the aviary was covered with sand, allowing daily removal of faecal matter and food leftovers.

The kea were fed three times a day with a balanced diet of fresh vegetables, fruits, dairy products, seeds and meat. Fresh drinking water was available ad libitum. Veterinary stock control was done by the Kleintierklinik at the University of Veterinary Medicine Vienna.

3.3. Test subjects

11 out of the 22 (23 with the chick, that hatched during the time of testing) kea of the group started to participate in the touchscreen experiments (see Table 1). The other 11 birds were either inexperienced juveniles, adult birds that failed habituation attempts or birds, which despite several attempts to test them, did not show any interest in the task. Having finished the touchscreen task, all of the 11 birds started with the real objects task. However one kea, Coco, showed no interest in the new study design, and was therefore excluded from the sample after three attempts to test her. Another bird, Plume, refused to participate after 11 sessions for no obvious reason. All of the birds lived together in one social group during the time of testing, except for Anu. He was partly separated from the group, due to the treatment of the consequences of an infection of his partner Elvira. She lived temporarily in a small compartment, apart from the big aviary.

Table 1: Individuals that were tested; prior experience with the touchscreen is indicated by listing the concerning studies: O'Hara et al. 2012 (O12), O'Hara et al. 2015 (O15), Wein et al. 2015 (W15)

Name	**Sex**	**Hatched**	**Parent/ hand reared**	**Experience with touchscreen**	**Participation ended**
Anu (An)	M	2007	hand	O15, W15	-
Coco (Co)	F	2007	hand	O15, W15	Beginning of real obj. (Aug. 1st)
Kermit (Ke)	M	2004	hand	O12, O15, W15	-
Lilly (Ly)	F	2007	hand	O15, W15	-
Papu (Pu)	F	2013	parent/hand	-	-
Paul (Pa)	M	2010	parent	W15	-
Pick (Pi)	M	2004	hand	O12, O15, W15	-
Plume (Pl)	F	2007	hand	O12, O15	During real obj. (Sept. 13th)
Roku (Ro)	M	2008	parent	O15, W15	-
Sunny (Sy)	F	2007	hand	O15, W15	-
Willy (Wy)	F	2007	hand	O12, O15, W15	-

4. Touchscreen experiment

Artificial stimuli, such as pictures and photographs, have been widely used in animal research, as they provide some advantages over solid objects. They are easier to handle, can be controlled easily in terms of timing and location and can be presented over and over again to the same or different subjects without much effort (D'Eath 1998). A central topic when using pictorial stimuli as representations of real objects is, however, the ability of the test subjects to perceive pictures indeed as representations of objects (Bovet and Vauclair 2000). A study investigating this question in kea showed that kea have the ability to transfer from objects to pictures (Wein et al. 2015).

The objects used in our experiment were all well known by the tested kea, as they were confronted with them on a daily basis. This familiarity and experience is thought to enhance performance in discrimination and recognition tasks with pictures (Bovet and Vauclair 2000). Some researchers suspect that birds might lack the ability to retrieve three-dimensional representations from two-dimensional pictures, when they have not had picture-object training (Soto and Wasserman 2010a) or prior exposure to pictures in general (Bovet and Vauclair 2000). However most of the birds in our study have had prior experience with picture-object recognition tasks or tasks involving pictures as representations of objects (O'Hara et al. 2015; Wein et al. 2015). If Soto and Wasserman's suspicion is true, then the experienced birds (see list in chapter 'Test subjects') should perform better than the inexperienced (Papu).

Soto and Wasserman (2010b) further suggest that photographs of real objects "more closely resemble the stimuli that are encountered by biological systems in the real world than the more commonly used artificial stimuli of the laboratory". Benefits of using a touchscreen are that this method enhances "data collection efficiency, reduces or eliminates possible experimenter bias and may aid interspecies comparability of tasks" (O'Hara et al. 2015). Wein et al. (2015) showed that kea have the ability for picture-object recognition, which should especially work with the pictures we used, as they depicted objects that the kea were confronted with on a daily basis. Also the sounds that were played as samples from a loudspeaker, which can be assumed to be perceived by the subjects as realistic as the real sounds they represent (Sturdy and Weisman 2006; Weisman and Spetch 2010), should enhance the performance, as they add additional cues, presumably leading to better recognition performance, when considering the abovementioned increased behavioural responses to multimodal stimuli.

4.1. Apparatus

The touchscreen we used is a combination of a 15-inch XGA colour TFT computer screen (Model G150XG01 by AU Optronics Corp., Taiwan) with a display area of 304 mm * 228 mm (381 mm diagonal) and a resolution of 1,024 * 768 pixels, and a 15-inch IR 'CarrollTouch' touchframe (Model D87587-001 by Elo, Menlo Park, CA, USA) for detecting the subjects' responses. Connected with this screen was a CPU and a modified operant conditioning system with an automatic feeder that distributes rewards directly after a correct response. We used a portable loudspeaker (Model GO by JBL, LA, USA), which was connected with the CPU. The program used for testing was CognitionLab (Version 1.9), which controlled the presentation of the stimuli, the reward system, and also recorded the responses (Steurer et al. 2012).

To interact with the touchscreen the subjects had to enter a wooden cabin in the experimental compartment of the aviary. This cabin was enclosed by walls on three sides, leaving one side open for the birds to enter (or leave) voluntarily. The whole construction was introduced to this kea group in 2009 (O'Hara et al. 2012) and has since been used for several experiments (e.g O'Hara et al. 2015, 2016; Ravignani et al. 2015; Wein et al. 2015).

4.2. Stimuli

The stimuli we used were photographs and sound recordings of objects that the kea encounter on a daily basis and are familiar with. The four objects we started the training with (see Fig. 4) were a food bowl, a sliding door (entrance to aviary), a shovel with a rake, and a water tap from one of the ponds. Four more objects (brush, drill, stones, opaque sliding wall) were initially planned for further training steps (transfer), but none of the kea reached the criterion to finish the first phase, so these were not used.


bowl door rake water

Figure 4: Stimuli for the touchscreen experiment, which were pictured in the context where they make the sound that the birds were required to use to discriminate them by: feeding bowl being banged on the feeding table, sliding door shut, rake being used to gather dirt from aviary floor, water from spout filling pond.

Five pictures of each object were used, each depicting it from a slightly different perspective. The photographs of the bowl and the rake included hands, as the objects are only known to the kea to produce sounds when being handled by a human. For each kea two of the objects were randomly selected for the first training phase and would remain the same throughout all sessions.

The photographs were taken with a digital single-lens reflex camera (Canon EOS 1000d 10.1 megapixel, Canon Inc. Tokyo, Japan) in the aviary in natural light, where the objects are normally perceived by the kea. The sounds were recorded with a directional Sennheiser ME66 microphone, connected with a Zoom H4n handheld recording device and edited with the Audacity 2.1.2. audio software (released 2016, http://www.audacityteam.org/). For each object the best sound recording was chosen, to be the sample for all five pictures.

4.3. Procedure

Each session consisted of 20 trials and in each trial photographs of two different objects were presented simultaneously. The stimulus pairs and their order over the course of the session were semi-randomized with a ‘random generator’, set up in Excel 2013. The stimuli were located horizontally in the middle of the screen, one third and two thirds from the screen’s side frame. The side of the S+ and the S- were randomized by the program (CognitionLab).

The S+ was always the picture of the object that was associated with the (approximately one second long) sound sample that was played once, as soon as the pictures appeared (, after a centred trigger-stimulus, a white square, was pecked on). Sample and stimuli were thus presented at the same time, making the task easier, as no auditory working memory (, which was not part of the research question) would be required for it (Hashiya and Kojima 2001). The inter-trial-interval was set to one second and consisted of a black screen.

Each object served as the S+, respectively the S-, for 10 of the 20 trials. Every bird was confronted with 2 different objects, each of which represented though 5 photographs, resulting in 10 different pictures appearing in each session. Every picture was thus twice a correct choice and twice an incorrect choice. There were never two pictures of the same object in one trial.

When a subject touched the S+ a reward (an eighth of a peanut seed) was delivered through the automatic feeding system to the reward tray, below the touchscreen. Pecking the S- led to correction trials. After approximately 15 sessions we stopped allowing correction trials, as none of the subjects had reached results above chance level. We assumed stopping correction trials might enhance performance, as with them an incorrect choice was in no way disadvantageous for the subject; it would get 20 rewards per session regardless of the performance. This change in experimental design, however, did not change the outcome. So we introduced another change after approximately 30 sessions. Before that the S+ and S- were semi-randomized over the course of the 20 trials. The new setup then had the same S+ and S- for the first 10 trials, which would then switch for the second 10 trials (sides, however, were still randomized by the program).

4.4. Data collection

The data was recorded automatically by the CognitionLab program and included the correct/ incorrect first choices, correct/ incorrect choices in correction trials, the date, the time, and the position of the stimuli. Additionally the experimenter recorded date, time, and number of correct first choices and correction trials for each session manually.

4.5. Statistical analysis & results

After over 40 sessions with every individual still performing at chance level and several birds showing signs of side biases we decided to check for these observed tendencies statistically, to determine how to proceed. The statistical tests were performed with the data analysis program IBM SPSS Statistics 23 (2015) for Windows. Two-tailed binomial tests, with the test proportion set to 0.5, were used to see if any of the subjects performed above chance level and if the side biases were significant.

The test was conducted with the results from the last two sessions of each subject. On the group-level 225 out of 440 choices were correct. This result is not significant (binomial test: N=440, p=0.668). To get a more comprehensive picture and to also check for side biases on an individual level binomial tests were made for every bird.

The p-values of the correct choices and the chosen sides over the last two sessions were calculated for each kea (see Table 2). Confirmation of the null hypothesis (= performance is at chance level, expected proportion: 0.5) would for the choices mean that the birds did not significantly choose the correct stimulus, for the side it would mean that no side bias has developed.

Table 2: p-values for each individual. Only Kermit and Lilly had significant choices. All kea except for Kermit, Papu and Willy significantly chose one side more often than the other, hinting at side biases

Individual	**Stimulus choices**	**Side choices**
Anu	0,636	<0,001
Coco	0,268	0,002
Kermit	**0,039** (incorrect choices)	**0,636**
Lilly	**0,039** (correct choices)	<0,001
Paul	1,000	0,016
Pick	0,875	<0,001
Plume	0,875	<0,001
Papu	0,875	**0,154**
Roku	0,430	<0,001
Sunny	0,154	<0,001
Willy	0,636	**0,081**

4.6. Discussion 1

4.6.1. Results

Most birds chose the stimuli seemingly randomly, while having established side biases (see Table 2). Only two kea (Kermit, Lilly) seemed to perform significantly above chance level ($p<0.05$). Kermit, however, significantly chose the incorrect stimulus, while Lilly seems to have established a side bias. Only two birds, other than Kermit, seem to have not established a side bias ($p>0.05$), but none of them was significant in their choices. None of the birds therefore seem to have learnt to solve the task properly.

The development of a side bias in several individuals can be viewed as a “last resort strategy”, which is not the cause, but the symptom of a failure to solve a task properly (Huber et al. 2013), resulting in the tendency to choose the stimuli on one side more often than on the other. This lower-level behaviour often interferes with more elaborate strategies in cognitive tasks (Ravignani et al. 2015).

We actually expected it to take some time for the kea to learn the matching-to-sample task on the touchscreen, as O’Hara et al. (2015) already showed that it takes longer for them to learn a discrimination on this device, compared to the same task with real objects. We did not, however, expect that after over 40 sessions (over 800 trials) not one of the birds would reach performance above chance level.

Due to the evident failure of our test subjects to perform successfully we decided to stop the experiments within the touchscreen design, as we expected further sessions just to strengthen the side biases, instead of actual progress. However, if the kea were unable to use the pictures as representations of the depicted objects and use the sound cues to match them, the question arises why.

4.6.2. Possible reasons why the touchscreen experiment failed

Anatomical, physiological and behavioural investigations hint at substantial differences between human and avian visual and auditory perception (Delius et al. 1999). All of the used devices are designed for human perception, so what will be discussed are possible methodological problems with the experimental setup that might have interfered with the kea's performance.

4.6.2.1. Sound stimuli

Martinez and Matsuzawa (2009a; b) hint at that the possibility that in experiments with animals, recorded acoustic stimuli might not be experienced by the subjects as similar as natural sounds. So for our experimental design it could be questionable if the kea were actually able to perceive the recorded object sounds as equivalent to those produced in a real situation. Sturdy and Weisman (2006) by contrast note that recorded sounds played from speakers "evoke much the same responses as their naturally occurring counterparts". Studies with kea indeed showed that they respond to audio recordings of conspecifics in the same way as to real-life vocalizations (Schwing et al. 2017), implying that our recorded object sounds should also be perceived by them as realistic as real ones. Considering this we can rule out the possibility that the sounds per se were not perceivable in an adequate manner.

Another topic is that of location of the stimulus sources. Harrison and Pratt (1977) and Harrison (1984) stress that in auditory choice-tasks with animals the response site should be right next to the sound source, facilitating fast learning, in contrast to separate spatial locations, where considerably more trials are needed for acquisition. Also Lickliter and Bahrick (2000) emphasize the importance of spatial colocation of auditory stimulus and visual cue in intersensory perception tasks, which should enhance the performance. Taking these considerations into account one has to conclude that the setup for our touchscreen task was not optimal, as the visual stimuli on the touchscreen were separate from the sound source (loudspeaker behind wooden wall, above the touchscreen). Now one could argue that to test CMP the sound should not come from S+, as then not the association between a picture and a sound is tested, but sound localization. Pairing of the stimuli, however, would have been a good training step to introduce the subjects to an experimental setup that incorporates visual and auditory stimuli.

4.6.2.2. Picture stimuli

To understand possible problems with our picture stimuli we have to consider the differences between human and avian visual perception, as the technology used to present them (photographs, screen) is produced to fit the requirements of our trichromatic visual system, not the avian one (Fagot and Parron 2010; Weisman and Spetch 2010). The most apparent difference is the amount of cones and their wavelength-sensitivities. Birds have at least four types of colour receptors in their retina, compared to three in humans, which can furthermore be assumed to have different maxima of wavelength-sensitivity, leading to a qualitatively different colour perception (Cuthill et al. 2000; Fagot and Parron 2010). The additional cone type of birds enables them to perceive UV wavelengths (Bennett and Cuthill 1994; Eaton 2005), but this UV component of light is absent in screens, possibly making depictions on them even harder to recognize for birds. These differences possibly caused the kea in our experiment to perceive the coloured photographs in a way, not comparable to what we humans see on them.

But not only colour perception is quite different (and in some respects superior) in birds compared to humans, so is visual acuity, which is higher in birds (D'Eath 1998). Combined with the subjects' possibility to inspect the pictures closely in our experiment this might have revealed the pixels on the screen (Dawkins and Woodington 1997) and disturb proper perception of the depicted objects. The outstanding visual abilities of birds when it comes to colour perception and visual acuity might therefore, paradoxically, have impaired the kea's performance in the touchscreen task.

Another important aspect of picture-object recognition are cues that enable to see three-dimensionality in pictures (Weisman and Spetch 2010). Especially depth cues play a role here, but many get lost through pictures, such as those through stereoscopic vision, motion parallax or focusing cues (D'Eath 1998; Bovet and Vauclair 2000). Additionally our test subjects were able to move relative to the screen, which could have led to a distortion of the depicted object. This might be an even bigger problem, as the stimuli were photographs and not just shapes (D'Eath 1998). What could also have been a problem is the change in size of the objects' representation on the pictures, compared to the real objects (D'Eath 1998; Bovet and Vauclair 2000). These general problems of picture perception, combined with the abovementioned potential problems of avian vision, could have in sum impaired the kea's perception of the depicted objects.

Despite these potential obstacles to proper picture perception it has also to be noted that the problems cannot be as drastic as it might seem, as kea have performed successfully on several touchscreen tasks. If the difficulties were indeed grounded in the impaired visual information then the results of O'Hara et al. (2015), Ravignani et al. (2015) and Wein et al. (2015) would have to be explained. A possible reason for their successful use of visual stimuli on screens might be the difference of the presentation of the target object. While the objects were visually embedded in their environment in our pictures, the successful touchscreen experiments used as stimuli either abstract computer generated elements (Ravignani et al. 2015) or photographs that have been edited (O'Hara et al. 2015; Wein et al. 2015). In these cases the stimuli could be discriminated not only on the basis of colour or three-dimensional perception, but also based on their shapes as they were presented in front of neutral (e.g. black) backgrounds, isolated from potentially disturbing surroundings and therefore increasing contrast between object and background. Our stimuli did not promote such a strategy, as the background was not edited and showed a 'natural setting' (kea aviary), where the objects are seen by the kea normally. Now one could argue that this setting should enhance performance, but one could conclude on the contrary that the complex visual surrounding, combined with altered colour perception or one of the other addressed possible perceptual problems, might confuse proper recognition of the depicted objects. Likewise D'Eath (1998) argues that transfer between simple objects and pictures of them is easier, than that of more complex stimuli. Bovet and Vauclair (2000) make similar considerations, when stating that stimuli that can be recognized due to shapes or silhouettes might be easier to perceive than more complex photographs.

4.6.2.3. Final remarks

Finally we should, however, also consider cognitive constraints that might have interfered with the kea’s perception and performance in the task. After all it is not easy to perceive a visual or auditory substitute as representing a real-life object. For this to function properly a dual representation, of the symbol itself as well as its relation to the referent, is necessary (Huber et al. 2013). Differences between species in their ability to recognize representations can therefore be based on perceptual, but also on cognitive processes, that are responsible for interpreting the sensory information (Spetch and Friedman 2006). Also, the failure could hint at the kea’s inability to perceive cross-modally, which we, however, do not expect. To investigate this, another experiment, targeting CMP, was conducted. As it was not clear if the failure of the kea to solve the task was based on specific perceptual problems in a single modality, or on the linking of the visual and auditory stimuli, the experimental setup was changed thoroughly to rule out as many potential problems as possible.

5. Object experiment

To overcome the described problems of the touchscreen task the experimental design was thoroughly changed to test with newly introduced, solid objects. This drastic step was taken to prevent a continuation of the errors of the kea in the previous task. Also new stimuli were chosen to rule out the possibility of object preferences influencing the performance (e.g. the food bowl, as it is associated with food) and also for reasons of feasibility (the sliding door or water tap cannot be moved to the testing compartment). Consequently two new objects (a wooden 'cube' and a wooden 'pyramid') were introduced, which were associated with different sounds (a 1000 Hz sound, and a 4000 Hz sound, respectively).

This new stimuli were chosen to overcome the obstacles of the touchscreen task. Most importantly all potential difficulties with using photographs on a computer screen were avoided. Using solid objects should enable the test subjects to make use of undistorted perception of colour and three-dimensionality, and therefore increase the discriminability of the stimuli (O'Hara et al. 2015).

For the training the sound was emitted directly from the object it should be associated with. This spatial pairing of visual and auditory stimuli is important to build associations of multimodal cues, as this is thought to happen only when some kind of commonality (e.g. spatial or temporal) between the stimuli is given (Calvert et al. 1998). In addition the reward should also not be separated from the stimulus, as it was in the touchscreen experiment. Miller and Murphy (1964) already showed that discrimination tasks are learnt faster, when the reward is spatially close to the S+. In the new experimental setup the reward (1/8 of a peanut seed) was placed under a small metal cup in front of the S+. This spatial pairing of object, sound, and reward also makes sense from an evolutionary point of view. After all, animals will have evolved towards associating objects and sounds that derive from the same location, where also food could be found. All sound producing objects in nature are spatially connected with the sound source (Harrison and Pratt 1977; Harrison 1984).

5.1. Stimuli


Figure 5: A kea with the 'pyramid' stimulus

Two stimuli were built for this experiment, wooden boxes (made from 15 mm thick chipboard) with different shapes. The 'pyramid' (see Fig. 5) consisted of an isosceles triangle (base: 225 mm, height: 212 mm) in the front and wooden boards on the sides (150 mm wide), the 'cube' (see Fig. 6) of a rectangle (150 mm * 183 mm) in the front, also with wooden boards on the sides (165 mm wide).

The boxes were not painted in different colours, as kea are known to show preferences for certain colours (e.g. yellow) (Auersperg et al. 2014), possibly resulting in preferring one object based on the colour. The background and the platform board, however, were painted blue with a nontoxic colour, to increase the contrast between stimulus and platform (344 mm * 315 mm)/ background (344 mm * 330 mm).


Figure 6: A kea with the 'cube' stimulus

Both objects had 23 identically arranged holes in the front for the auditory stimuli, coming from the inside of the boxes, to be perceived better (the sounds were initially played from JBL GO loudspeakers, connected with a Samsung Galaxy S4 Smartphone, but as the sounds seemed to be more audible when being played directly from the Smartphone we decided to not use the JBL loudspeakers for further testing). Each object was presented with a certain sound, whenever it was the S+ in a trial. The 'cube' was associated with a 1000 Hz tone, the 'pyramid' with a 4000 Hz tone (both created with Audacity 2.1.2.). In each of the trials one of the sounds was played continuously from the corresponding object (5 second sample, with 1 second break), making the task a simultaneous matching-to-sample task.

The frequency of the sounds was chosen to fit the kea's hearing range. Their highest auditory sensitivity is between 1000 and 4000 Hz, with a peak at 3000 Hz (Schwing et al. 2016). The frequencies chosen therefore represent the biggest difference within this window, implying the best possible discriminability, while ensuring optimal detectability. To make the acquisition of the audio-visual associations easier the objects were placed 160 cm from each other. This makes the choice of the subjects more decisive and enables for better learning of the differences, as it becomes clearer that the sound sources are separate (Harrison 1984).

5.2. Procedure

The training consisted of three consecutive phases, each of which was considered finished as soon as the subject had 18 (or more) correct choices out of 20 (≥ 90%) in two consecutive sessions. The tasks were two-choice discrimination tasks, in which always the object from which a sound was being played was the S+.

In the first phase we introduced the kea to the two objects between which they had to discriminate. For this in all 20 trials the same object was the S+, meaning that always the same sound was being played as an auditory cue. The sides were semi-randomized so that the same object would never be on the same side for more than two consecutive sessions. The one object (of the two possible) to start with as the S+ was pseudo-randomly assigned to each test subject, so that both starting groups would be the same size (starting with ‘pyramid’: An, Ly, Pa, Pi, Pu; with ‘cube’: Ke, Pl, Ro, Sy, Wy; Co refused to be tested after changing the experimental setup from the touchscreen).

In the second phase we introduced the kea to a change of the S+ within a session, meaning they were now also introduced to the sound of the other object. In the first 10 trials the same stimulus was S+ as in the previous phase, but would then be reversed to be the S- for the following 10 sessions. The sides were again semi-randomized, as in the previous phase. Up to here we did not know if the kea used the sound cue at all, it would be sufficient to solely learn the object being rewarded (with a midsession reversal in the second phase).

In the third phase we introduced the kea to having to choose the visual stimulus on the basis of the location of the sound cue. Not only the sides were randomized, but also which of the objects would be the S+ was semi-randomized in a way that each stimulus and each side would be correct for 50% of the trials. No side or S+ was the same for more than four consecutive trials. It was only with this step that we were able to show that the kea indeed used the sound cue for solving the task.

The stimuli were presented on the sandy aviary ground, with a distance of 160 cm between the objects and the midpoint between them being 160 cm away from the entrance to the testing compartment. The subjects were allowed to enter and solve the task immediately after the sliding door was opened. Each correct choice was rewarded with an eighth of a peanut seed, below a metal cup in front of the S+, a trial with an incorrect choice was not rewarded and not repeated. After each choice the kea was directed to return to the waiting compartment, which all of the subjects have learnt after a few sessions, so that they would return immediately after their choice. Between trials the kea had to wait in the neighbouring waiting compartment, while the experimenter arranged the stimuli. To rule out possible information being transported by preparing of the setup the movements of the experimenter were the same for the preparation of each trial (lift left reward cup to put in or remove reward or do the same movements without exchanging reward, do the same with the right reward cup, and change object location if it is different from previous trial). To prevent a Clever Hans effect the experimenter moved away from the stimuli and waited in a partly hidden place to the left of the setup, waiting there until the subject had made a choice and returned to the waiting compartment. Each of the birds was trained until it had finished all three of the described phases successfully.

5.3. Data collection

Results of the trials were recorded manually by the experimenter. The information noted were correct/ incorrect choices, date, time and unexpected occurrences, such as hints of side biases, kea ignoring the task or also observing what looked like a kea really listening closely to the sound cue.

5.4. Statistical analysis & results

All statistical tests were made with the data analysis program IBM SPSS Statistics 23 (2015) for Windows. Non-parametric tests were used to see if the subjects performed significantly above chance level. To check if every individual had learnt to significantly choose the correct stimulus a two-tailed binomial test was used, with the test proportion set to 0.5. The test was conducted with the results from the last two sessions of each bird in phase 3, thus after reaching the criterion. On the group-level 342 out of 360 choices were correct. This result is highly significant (binomial test: N=360, p<0.001). Also, every individual chose the correct stimulus accurately over the last two sessions, with the ‘worst’ birds being correct in 37 out of 40 trials (binomial test: N=40, p<0.001). Another aspect we were interested in was the amount of sessions needed to reach the criterion for each phase (see Fig. 7 and Fig. 8).

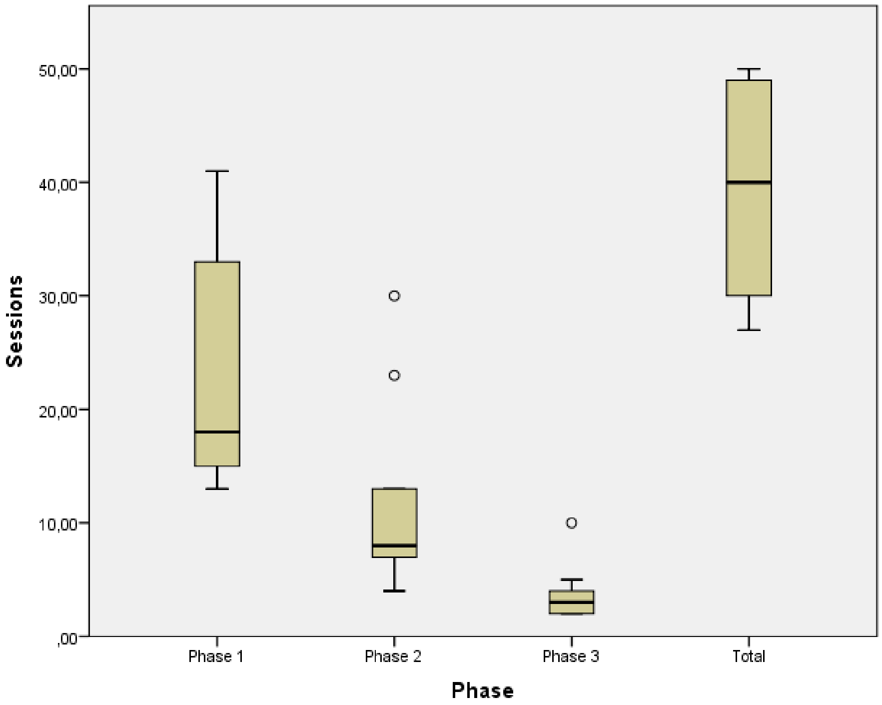

Figure 7: Number of sessions to reach the criterion for all three phases and the total number of sessions needed to reach final criterion. Performance improved with each phase, resulting in an average of 40 sessions, after which the final criterion was reached by the test subjects.

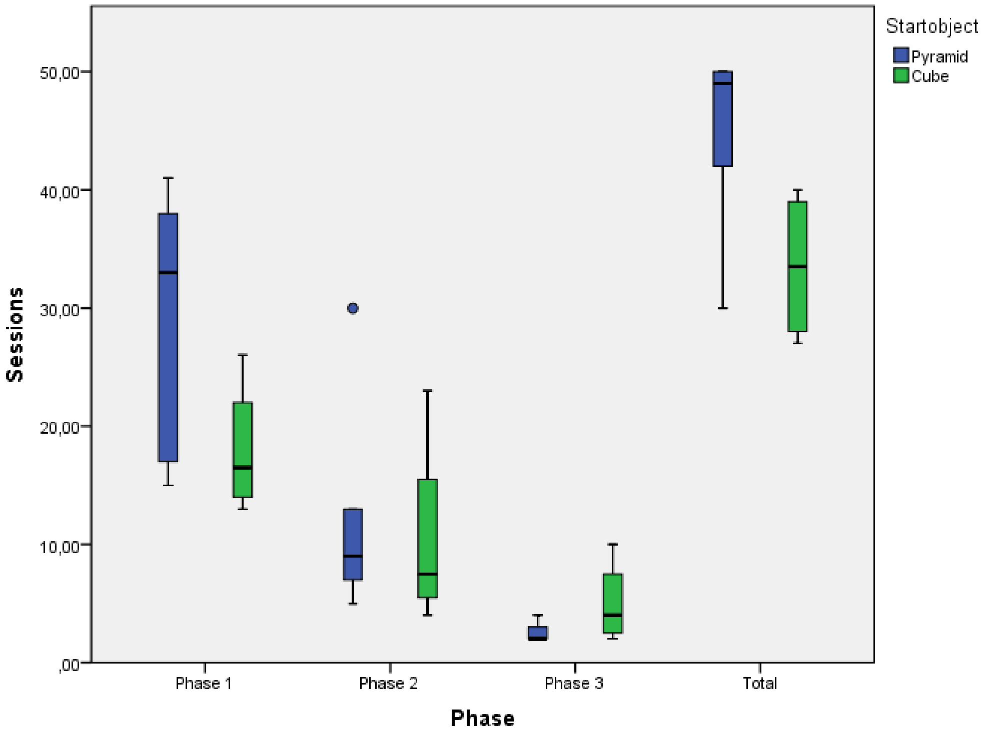

Figure 8: Number of sessions to reach criterion for all three phases and the total number of sessions needed to reach final criterion, split into group that started with 'pyramid' and group that started with 'cube'. Although implied by the boxplot the differences between groups were not significant.

The boxplot in Fig. 8 implies differences between the ‘starting-groups’. The birds that started with the ‘cube’ seemed to reach the criterion faster (especially in phase 1). To compare the groups and see if the differences are significant a Mann-Whitney-U-Test was performed. The results show, however, that none of the differences, neither of the phases nor of the total amount of sessions, is significant (see Appendix for a visual representation of the results of the individual subjects).

5.5. Discussion 2

5.5.1. Results of the object experiment

The results show that the kea have learnt to successfully use the sound cue to choose the correct object. The first and the second training step could have also been solved through other strategies, such as only paying attention to the visual appearance of the object, to the sound or to a combination of both, but only through the third phase we could definitely show that the kea indeed use the sound to solve the task.

In the first condition it would be sufficient to only pay attention to the visual appearance of the objects and learn to choose the stimulus that is rewarded. In the second phase (midsession reversal of S+ and S-) results could also be attained through a win-stay/lose-shift strategy, continuing to choose the same object, and as soon as it is not rewarded anymore switch to the other object. This was, however, not what we observed (see Appendix). After the switch from the first to the second phase some birds (Anu, Pick and Kermit) stuck to the wrong object after the mid of the session, staying with the formerly rewarded stimulus. But soon that tactic disappeared, leaving only the possibility to use the time into the session or to estimate the number of already finished trials to anticipate the switch to the other stimulus, or otherwise to pay attention to the sound and where it was coming from. Other birds even initially after the switch had more correct trials than would be expected when sticking to the formerly conditioned object (see Appendix: Papu, Roku and Willy had in the first session after the switch 14 correct; Paul and Sunny had 15 correct; Lilly even 16). These birds may have payed attention to the sound more than the others, but none seem to have learned in the first phase to use the sound direction as the only decisive cue, which would result in a performance as good as in the last sessions of the first condition. In fact this possible focus on only one modality is not what we wanted to accomplish, so the observed learning progress seems promising for generating a cross-modal association between visual and auditory information.

To investigate the role of the sound cue in the third and final condition also the object, which would be the S+ (and therefore emitting a sound), was semi-randomized over the course of the sessions. This would rule out that other strategies, which do not include using the sound, such as using knowledge about the outcome of previous trials to predict the future outcome, could be successful. Indeed, most of the birds seem to have learned to use the sound as a decisive cue in the first two conditions, as the third phase was finished very fast, often after only two sessions (as was the case with Anu, Lilly, Paul, Pick and Kermit). Interestingly, some kea that finished the second condition relatively fast took the longest to reach the criterion in the third step (Sunny: 5 sessions; Willy: 10 sessions). Especially Willy is interesting, as she was the first subject to reach the final criterion. Maybe these individuals had too few previous sessions to internalize to use the sound and were confused with the changing S+ and S- in the beginning. It has to be noted, however, that also with these birds the performance in the third phase was quite good (Sunny: always 16 correct trials per session or more; Willy: also always 16 correct or more, with one exception). As in the end all kea reached the final criterion we could show that they indeed do pay attention to the sound as a discriminating cue, at least to the direction it is coming from. This is a replication of the finding of Schwing et al. (2016) that kea are able to use a sound source as a discriminating cue, but now with a bigger sample size. In the mentioned study only three birds were tested. Interestingly these successfully tested kea were Anu, Frowin and Roku, two of which were also tested in our experiment and were amongst the three fastest individuals to reach the final criterion (see Appendix). This could be due to a special auditory ability in these individuals or maybe also be caused by their prior experience with an auditory task.

When considering the three conditions separately from another it becomes apparent that on the group level each step was finished faster than the previous one. This might reflect a continuous learning process of increasingly incorporating the sound cue in the decision making. It shows that for most of the birds the first step was the most difficult, but after they understood the concept they were quite fast at solving the second task and even faster at the third.

As discussed, reaching the third criterion showed that the birds have used the sound as a decisive cue. What, however, cannot be answered at this point is if the subjects payed attention to the sounds themselves and discriminated them based on their different frequencies (leading them to the associated object), or only to the direction of a sound source.

Indeed there are indications that they used the direction of the sound, as it was observed in several individuals in some trials that they moved towards one object, then stopped and listened, to finally choose the other object that emitted the sound. This might be because auditory spatial perception is more difficult from the starting position, as both objects are equidistant from there, than it is when already moving towards one object. This strategy suggests that the kea had used the directional cue in the first place, and not already the cross-modal association between sound and object. This step of aligning object and sound source is, however, very important for building cross-modal associations, as has been considered in the discussion for the touchscreen experiment above. The next logical step in testing the birds would be to separate sound source and object again, which would test if they have built cross-modal associations between them.

5.5.2. Future research

Due to time restrictions, it was not possible to continue research to the point where cross-modal association was tested, but with the results so far, we can conclude that kea do pay attention not only to visual, but also to auditory cues, when confronted with a cross-modal experiment. This, however, can only be the starting point in investigating CMP in this species. The obvious next step will be to separate the sound source and the visual stimulus again, after the subjects could learn the audio-visual association in the training so far. The experimental setup can stay quite similar to the one used for this thesis, but the sound should then come not from the S+, but from a 'neutral' location, optimally from the middle between the stimuli. If the birds have only learnt to move towards the sound source, they would then approach the loudspeaker in the middle. If they have built associations between the sounds and the objects, they should choose the S+.

It would also be possible to return to the touchscreen setup and use the objects and sounds introduced in the 'real objects task'. Discrimination should be easier, compared to the photographs used in the original touchscreen experiment, possibly overcoming the obstacles considered in 'Discussion 1'. Nevertheless, I would suggest continuing to work with real objects, as this experimental setup seems promising. Also we know little about the kea's vision, but we for sure can say that using pictures on a screen will not make the task easier (O'Hara et al. 2015).

Once CMP in kea has then successfully been shown, further research can move into several directions. One possibility is to investigate cross-modal individual recognition of con- or heterospecifics. Conspecific recognition is widespread in the animal kingdom (Leopold and Rhodes 2010), also when tested cross-modally (Kojima et al. 2003; Bovet and Deputte 2009; Proops et al. 2009; Adachi and Hampton 2011; Kondo et al. 2012; Kulahci et al. 2014). Some studies also suggest cross-modal recognition of heterospecifics in several species (Adachi and Fujita 2007; Adachi et al. 2007, 2009; Sliwa et al. 2011; Lampe and Andre 2012a; Proops and McComb 2012). This has so far only been tested with recognition of humans. The precondition for it seems to be extensive contact with individuals of the other species. The adaptive value for this ability might be the recognition of predators (Huber et al. 2013), but we know very little about this phenomenon so far. Also, one needs to be careful with terminology in this area of research. When offering tasks for cross-modal recognition, the subjects might be successful simply for the fact that they regard one stimuli as more familiar compared to another one (Tibbetts and Dale 2007; Huber et al. 2013). Should kea be successful at a task investigating cross-modal recognition of humans, then the question arises on what components this recognition is based. It would be interesting to test different visual displays, such as whole body, only face, body without face or maybe even just clothing.

Although the research question posed could not be answered I still regard the experiments conducted for this thesis to be a success. Possible experimental setups to investigate CMP in kea have been proposed and also applied, showing advantages and also potential obstacles to different approaches. On the basis of the results future research can be built on. Hopefully the studies conducted help to induce further studies in this fascinating field of research.

6. References

Adachi I, Fujita K (2007) Cross-modal representation of human caretakers in squirrel monkeys. Behav Processes 74:27–32. doi: 10.1016/j.beproc.2006.09.004

Adachi I, Hampton RR (2011) Rhesus monkeys see who they hear: Spontaneous cross-modal memory for familiar conspecifics. PLoS One. doi: 10.1371/journal.pone.0023345

Adachi I, Kuwahata H, Fujita K (2007) Dogs recall their owner’s face upon hearing the owner's voice. Anim Cogn 10:17–21. doi: 10.1007/s10071-006-0025-8

Adachi I, Kuwahata H, Fujita K, et al (2009) Plasticity of ability to form cross-modal representations in infant Japanese macaques. Dev Sci 12:446–452. doi: 10.1111/j.1467-7687.2008.00780.x

Auersperg AMI, Gajdon GK, Huber L (2009) Kea (Nestor notabilis) consider spatial relationships between objects in the support problem. Biol Lett 5:455–458. doi: 10.1098/rsbl.2009.0114

Auersperg AMI, Gajdon GK, Huber L (2010) Kea, Nestor notabilis, produce dynamic relationships between objects in a second-order tool use task. Anim Behav 80:783–789. doi: 10.1016/j.anbehav.2010.08.007

Auersperg AMI, Huber L, Gajdon GK (2011) Navigating a tool end in a specific direction: stick-tool use in kea (Nestor notabilis). Biol Lett 7:825–8. doi: 10.1098/rsbl.2011.0388

Auersperg AMI, Oswald N, Domanegg M, et al (2014) Unrewarded Object Combinations in Captive Parrots. Anim Behav Cogn 1:470–488. doi: 10.12966/abc.11.05.2014

Bahrick LE, Lickliter R, Flom R (2004) Intersensory Redundancy Guides the Development of Selective Attention , Perception , and Cognition in Infancy. Curr Dir Psychol Sci 13:99–102. doi: 10.1111/j.0963-7214.2004.00283.x

Bennett ATD, Cuthill IC (1994) Ultraviolet Function ? Vision in Birds : What is its. J Comp Physiol 34:1471–1478. doi: 10.1016/0042-6989(94)90149-X

Bischof H-J, Engelage J (1985) Flash evoked responses in a song control nucleus in the zebra finch (Taeniopygia guttata castanotis). Brain Res 326:370–374. doi: 10.1016/0006-8993(85)90048-4

Boesch C (2007) What makes us human (Homo sapiens)? The challenge of cognitive cross-species comparison. J Comp Psychol 121:227–40. doi: 10.1037/0735-7036.121.3.227

Bovet D, Deputte BL (2009) Matching vocalizations to faces of familiar conspecifics in grey-cheeked mangabeys (Lophocebus albigena). Folia Primatol 80:220–232. doi: 10.1159/000235688

Bovet D, Vauclair J (2000) Picture recognition in animals and humans. Behav Brain Res 109:143–165. doi: 10.1016/S0166-4328(00)00146-7

Calvert GA, Brammer MJ, Iversen SD (1998) Crossmodal identification. Trends Cogn Sci 2:247–253. doi: 10.1016/S1364-6613(98)01189-9

Cowey A, Weiskrantz L (1975) Demonstration of cross-modal matching in rhesus monkeys, Macaca Mulatta. Neuropsychologia 13:117–120. doi: 10.1016/0028-3932(75)90057-3

Cuthill IC, Partridge JC, Bennett ATD, et al (2000) Ultraviolet Vision in Birds. Adv study Behav 29:159–214.

D’Eath RB (1998) Can video images imitate real stimuli in animal behaviour experiments? Biol Rev 73:267–292. doi: 10.1111/j.1469-185X.1998.tb00031.x

Davenport RK (1977) Cross-Modal Perception: A Basis for Language? In: Language Learning in a Chimpanzee: The Lana Project. pp 73–83

Davenport RK (1976) Cross-modal perception in apes. Ann N Y Acad Sci 280:143–149.

Davenport RK, Rogers CM, Russell IS (1975) Cross-modal perception in apes : altered visual cues and delay. Neuropsychologia 13:229–235.

Davenport RK, Rogers CM, Russell IS (1973) Cross modal perception in apes. Neuropsychologia 11:21–28. doi: 10.1016/0028-3932(73)90060-2

Dawkins MS, Woodington A (1997) Distance and the presentation of visual stimuli to birds. Anim Behav 54:1019–25. doi: 10.1006/anbe.1997.0519

Del Hoyo J, Elliot A, Sargatal J (1997) Handbook of the birds of the world, Vol. 4. Sandgrouse to Cuckoos.

Delius J, Emmerton J, Hörster W, et al (1999) Picture-object recognition in pigeons. Cah

Psychol Cogn 18:621–656.

Diamond J, Bond AB (1999) Kea, Bird of Paradox: The Evolution and Behavior of a New Zealand Parrot. University of California Press

Eaton MD (2005) Human vision fails to distinguish widespread sexual dichromatism among sexually “monochromatic” birds. Proc Natl Acad Sci U S A 102:10942–10946. doi: 10.1073/pnas.0501891102

Emery NJ (2006) Cognitive ornithology: the evolution of avian intelligence. Philos Trans R Soc Lond B Biol Sci 361:23–43. doi: 10.1098/rstb.2005.1736

Ernst MO, Bülthoff HH (2004) Merging the senses into a robust percept. Trends Cogn Sci 8:162–169. doi: 10.1016/j.tics.2004.02.002

Ettlinger G (1960) Cross-modal transfer of training in monkeys. Behaviour 16:56–65.

Ettlinger G (1973) The transfer of information between sense-modalities: A neuropsychological review. In: Memory and Transfer of Information. pp 43–64

Ettlinger G, Blakemore CB (1967) Cross-modal matching in the monkey. Neuropsychologia 5:147–154.

Ettlinger G, Wilson WA (1990) Cross-modal performance: behavioural processes, phylogenetic considerations and neural mechanisms. Behav Brain Res 40:169–192. doi: 10.1016/0166-4328(90)90075-P

Fagot J, Parron C (2010) Picture Perception in Birds: Perspective from Primatologists. Comp Cogn Behav Rev 5:132–135. doi: 10.3819/ccbr.2010.50007

Gsell AC, Hagelin JC, Brunton DH (2012) Olfactory sensitivity in Kea and Kaka. Emu 112:60–66. doi: 10.1071/MU11052

Guilford T, Dawkins MS (1991) Receiver Psychology and the Evolution of Animal Signals. Anim Behav 42:1–14. doi: 10.1016/s0003-3472(05)80600-1

Güntürkün O (2005) The avian “prefrontal cortex” and cognition. Curr Opin Neurobiol 15:686–693. doi: 10.1016/j.conb.2005.10.003

Harrison JM (1984) The functional analysis of auditory discrimination. J Acoust Soc Am

751845-1854 1848–1854.

Harrison JM, Pratt SR (1977) By Location of Auditory Stimuli : J Exp Anal Behav 3:243–251.

Hashiya K, Kojima S (2001) Acquisition of auditory-visual intermodal matching-to-sample by a chimpanzee (Pan troglodytes): Comparison with visual-visual intramodal matching. Anim Cogn 4:231–239. doi: 10.1007/s10071-001-0118-3

Huber L (2000) Wie das Neue in die Gehirne kommt. Emergenz und Chaos in neuronalen Prozessen. In: Huber L (ed) Wie das Neue in die Welt kommt. WUV Universitätsverlag, pp 157–174

Huber L (2007) Emulation learning: the integration of technical and social cognition. In: Nehaniv C, Dautenhahn K (eds) Imitation and social learning in robots, humans and animals : behavioural, social and communicative dimensions. pp 427–439

Huber L, Gajdon GK (2006) Technical intelligence in animals: The kea model. Anim Cogn 9:295–305. doi: 10.1007/s10071-006-0033-8

Huber L, Racca A, Scaf B, et al (2013) Discrimination of familiar human faces in dogs (Canis familiaris). Learn Motiv 44:258–269. doi: 10.1016/j.lmot.2013.04.005

Huber L, Rechberger S, Taborsky M (2001) Social learning affects object exploration and manipulation in keas, Nestor notabilis. Anim Behav 62:945–954. doi: 10.1006/anbe.2001.1822

Hultsch H, Schleuss F, Todt D (1999) Auditory-visual stimulus pairing enhances perceptual learning in a songbird. Anim Behav 58:143–149. doi: 10.1006/anbe.1999.1120

Iwaniuk AN, Dean KM, Nelson JE (2005) Interspecific allometry of the brain and brain regions in parrots (Psittaciformes): Comparisons with other birds and primates. Brain Behav Evol 65:40–59. doi: 10.1159/000081110

Jarvis ED, Güntürkün O, Bruce L, et al (2005) Avian brains and a new understanding of vertebrate brain evolution. Nat Rev Neurosci 6:151–159. doi: 10.1038/nrn1606

Johnstone RA (1996) Multiple displays in animal communication:`Backup signals’ and `multiple messages'. Philos Trans R Soc B Biol Sci 351:329. doi: 10.1098/rsta.1892.0001

King AJ, Calvert GA, Hospital JR (2001) Multisensory integration: Perceptual grouping by eye and ear for Functional Magnetic Resonance Imaging of the Brain (FMRIB). Curr Biol 11:322–325. doi: 10.1016/S0960-9822(01)00175-0

Knudsen EI, Brainard MS (1995) Creating a unified representation of visual and auditory space in the brain. Annu Rev Neurosci 18:19–43. doi: 10.1146/annurev.neuro.18.1.19

Kojima S, Izumi A, Ceugniet M (2003) Identification of vocalizers by pant hoots, pant grunts and screams in a chimpanzee. Primates 44:225–230. doi: 10.1007/s10329-002-0014-8

Kondo N, Izawa E-I, Watanabe S (2012) Crows cross-modally recognize group members but not non-group members. Proc R Soc B Biol Sci 279:1937–1942. doi: 10.1098/rspb.2011.2419

Kulahci IG, Drea CM, Rubenstein DI, Ghazanfar A a (2014) Individual recognition through olfactory-auditory matching in lemurs. Proc Biol Sci 281:20140071. doi: 10.1098/rspb.2014.0071

Lalanne C, Lorenceau J (2004) Crossmodal integration for perception and action. J Physiol Paris 98:265–279. doi: 10.1016/j.jphysparis.2004.06.001

Lampe JF, Andre J (2012a) Cross-modal recognition of human individuals in domestic horses (Equus caballus). Anim Cogn 15:623–630. doi: 10.1007/s10071-012-0490-1

Lampe JF, Andre J (2012b) Cross-modal recognition of human individuals in domestic horses (Equus caballus). Anim Cogn 15:623–630. doi: 10.1007/s10071-012-0490-1

Lefebvre L, Reader SM, Sol D (2004) Brains, innovations and evolution in birds and primates. Brain Behav Evol 63:233–246. doi: 10.1159/000076784

Leopold D a, Rhodes G (2010) A Comparative View of Face Perception. J Comp Psychol 124:233–251. doi: 10.1037/a0019460.A

Lickliter R, Bahrick LE (2000) The Development of Infant Intersensory Perception: Advantages of a Comparative Convergent-Operations Approach. Psychol Bull Aslin Smith 126:260–280. doi: 10.1037/0033-2909.126.2.260

Lickliter R, Bahrick LE (2004) Perceptual Development and the Origins of Multisensory Responsiveness. In: The Handbook of multisensory processes. pp 643–654

Malone DR, Tolan JC, Rogers CM (1980) Cross-modal matching of objects and photographs in the monkey. Neuropsychologia 18:693–697.

Martinez L, Matsuzawa T (2009a) Visual and auditory conditional position discrimination in chimpanzees (Pan troglodytes). Behav Processes 82:90–94. doi: 10.1016/j.beproc.2009.03.010

Martinez L, Matsuzawa T (2009b) Effect of species-specificity in auditory-visual intermodal matching in a chimpanzee (Pan troglodytes) and humans. Behav Processes 82:160–163. doi: 10.1016/j.beproc.2009.06.014

Meredith MA (2002) On the neuronal basis for multisensory convergence: A brief overview. Cogn Brain Res 14:31–40. doi: 10.1016/S0926-6410(02)00059-9

Miller RE, Murphy J V. (1964) Influence of the spatial relationships between the cue, reward, and response in discrimination learning.

O’Hara M, Gajdon GK, Huber L (2012) Kea Logics: How These Birds Solve Difficult Problems and Outsmart Researchers. In: Logic and Sensibility. pp 23–38

O’Hara M, Huber L, Gajdon GK (2015) The advantage of objects over images in discrimination and reversal learning by kea, Nestor notabilis. Anim Behav 101:51–60. doi: 10.1016/j.anbehav.2014.12.022

O’Hara M, Schwing R, Federspiel I, et al (2016) Reasoning by exclusion in the kea (Nestor notabilis). Anim Cogn 19:965–975. doi: 10.1007/s10071-016-0998-x

Over R, Mackintosh NJ (1969) Cross-modal transfer of intensity discrimination by rats. Nature 224:918–919. doi: 10.1038/224918a0

Partan SR, Marler P (1999) Communication goes multimodal. Science (80-) 283:1272–1273. doi: 10.1017/CBO9781107415324.004

Proops L, McComb K (2012) Cross-modal individual recognition in domestic horses (Equus caballus) extends to familiar humans.

Proops L, McComb K, Reby D (2009) Cross-modal individual recognition in domestic horses (Equus caballus). Proc Natl Acad Sci U S A 106:947–951. doi: 10.1073/pnas.0809127105

Ratcliffe VF, McComb K, Reby D (2014) Cross-modal discrimination of human gender by domestic dogs. Anim Behav 91:127–135. doi: 10.1016/j.anbehav.2014.03.009

Ravignani A, Westphal-Fitch G, Aust U, et al (2015) More than one way to see it: Individual heuristics in avian visual computation. Cognition 143:13–24. doi: 10.1016/j.cognition.2015.05.021

Romanski LM (2007) Representation and integration of auditory and visual stimuli in the primate ventral lateral prefrontal cortex. Cereb Cortex 17:61–69. doi: 10.1093/cercor/bhm099

Roskies AL (1999) Review. 24:7–9.

Rowe C (1999) Receiver psychology and evolution of multicomponent signals. Anim Behav 58:921–931.

Rowe C (2002) Sound improves visual discrimination learning in avian predators. Proc R Soc B Biol Sci 269:1353–1357. doi: 10.1098/rspb.2002.2012

Rowe C, Guilford T (1999) The evolution of multimodal warning displays. Evol Ecol 13:655–671. doi: 10.1023/A:1011021630244

Sakata S, Yamamori T, Sakurai Y (2004) Behavioral studies of auditory-visual spatial recognition and integration in rats. Exp Brain Res 159:409–417. doi: 10.1007/s00221-004-1962-6

Schiller P v. (1932) Intersensorielle Transposition bei Fischen. Z Vgl Physiol 19:304–309.

Schloegl C, Dierks A, Gajdon GK, et al (2009) What you see is what you get? Exclusion performances in ravens and keas. PLoS One. doi: 10.1371/journal.pone.0006368

Schwing R (2014) The Acoustic Signal Pathway of the Kea (Nestor norabilis). University of Auckland

Schwing R, Nelson XJ, Parsons S (2016) Audiogram of the kea parrot, Nestor notabilis. J Acoust Soc Am 140:3739–3744. doi: 10.1121/1.4967757

Schwing R, Nelson XJ, Wein A, Parsons S (2017) Positive emotional contagion in a New Zealand parrot. Curr Biol 27:R213–R214. doi: 10.1016/j.cub.2017.02.020

Shimizu T (2009) Why Can Birds Be So Smart? Background, Significance, and Implications of the Revised View of the Avian Brain. Comp Cogn Behav Rev 4:103–115. doi: 10.3819/ccbr.2009.40011

Sliwa J, Duhamel J-R, Pascalis O, Wirth S (2011) Spontaneous voice-face identity matching by rhesus monkeys for familiar conspecifics and humans. Proc Natl Acad Sci 108:1735–1740. doi: 10.1073/pnas.1008169108

Smith F V., Bird MW (1963) The relative attraction for the domestic chick of combinations of stimuli in different sensory modalities. Anim Behav 11:300–305. doi: 10.1016/S0003-3472(63)80115-3

Soto F a, Wasserman E a (2010a) Comparative Vision Science: Seeing Eye to Eye? Comp Cogn Behav Rev 5:148–154. doi: 10.3819/ccbr.2010.50011.Comparative

Soto FA, Wasserman EAE. (2010b) Error-driven learning in visual categorization and object recognition: a common-elements model. Psychol Rev 117:349–381. doi: 10.1037/a0018695

Spence C (2011) Crossmodal correspondences: A tutorial review. 971–995. doi: 10.3758/s13414-010-0073-7

Spetch ML, Friedman a. (2006) Comparative Cognition of Object Recognition. Comp Cogn Behav Rev 1:12–35. doi: 10.3819/ccbr.2008.10002

Steurer MM, Aust U, Huber L (2012) The Vienna comparative cognition technology (VCCT): An innovative operant conditioning system for various species and experimental procedures. Behav Res Methods 909–918. doi: 10.3758/s13428-012-0198-9

Stobbe N, Westphal-Fitch G, Aust U, Fitch WT (2012) Visual artificial grammar learning: comparative research on humans, kea (Nestor notabilis) and pigeons (Columba livia). Philos Trans R Soc B Biol Sci 367:1995–2006. doi: 10.1098/rstb.2012.0096

Sturdy CB, Weisman RG (2006) Rationale and methodology for testing auditory cognition in songbirds. Behav Processes 72:265–272. doi: 10.1016/j.beproc.2006.03.007

Tibbetts EA, Dale J (2007) Individual recognition: it is good to be different. Trends Ecol Evol 22:529–537. doi: 10.1016/j.tree.2007.09.001

Van Kampen HS, Bolhuis JJ (1991) Auditory Learning and Filial Imprinting in the Chick. Behaviour 117:303–319. doi: 10.1163/156853991X00607

Van Kampen HS, Bolhuis JJ (1993) Interaction between auditory and visual learning during filial imprining. Anim Behav 45:623–625. doi: 10.1006/anbe.1993.1074

Van Kampen HS, Bolhuis JJ (1992) An Evaluation of Auditory Learning in Filial Imprinting. Behaviour 122:195–230. doi: 10.1163/156853992X00516

Watanabe S, Huber L (2006) Animal logics: Decisions in the absence of human language. Anim Cogn 9:235–245. doi: 10.1007/s10071-006-0043-6

Wein A, Gajdon GK, Schwing R (2015) Picture - Object Recognition in Kea (Nestor notabilis). Ethology 121:1059–1070. doi: 10.1111/eth.12423

Weisman R, Spetch ML (2010) Determining When Birds Perceive Correspondence Between Pictures and Objects: A Critique. Comp Cogn Behav Rev 5:117–131. doi: 10.3819/ccbr.2010.50006

Werdenich D, Huber L (2006) A case of quick problem solving in birds: String pulling in keas, Nestor notabilis. Anim Behav 71:855–863. doi: 10.1016/j.anbehav.2005.06.018

Yehle AL, Ward JP (1970) Cross-modal transfer of a specific discrimination in the bushbaby (Galago senegalensis). J Comp Physiol Psychol 73:74–77. doi: 10.1037/h0029812

7. Appendix

The following graphs show the learning curves of all individuals, differently coloured for each training phase, with the number of correct trials per session mapped directly in the graph. The first five kea are from the group that started with the “pyramid” (An, Ly, Pa, Pi, Pu), the last four kea are from the “cube” group (Ke, Ro, Sy, Wy):

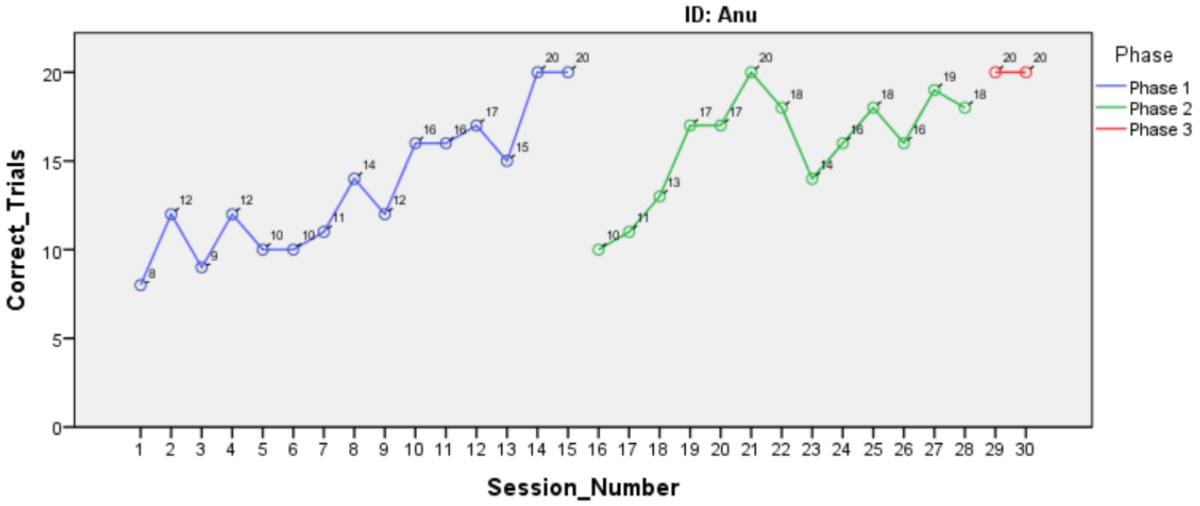

Figure 9: Anu's results, showing the amount of correct trials per session for each phase

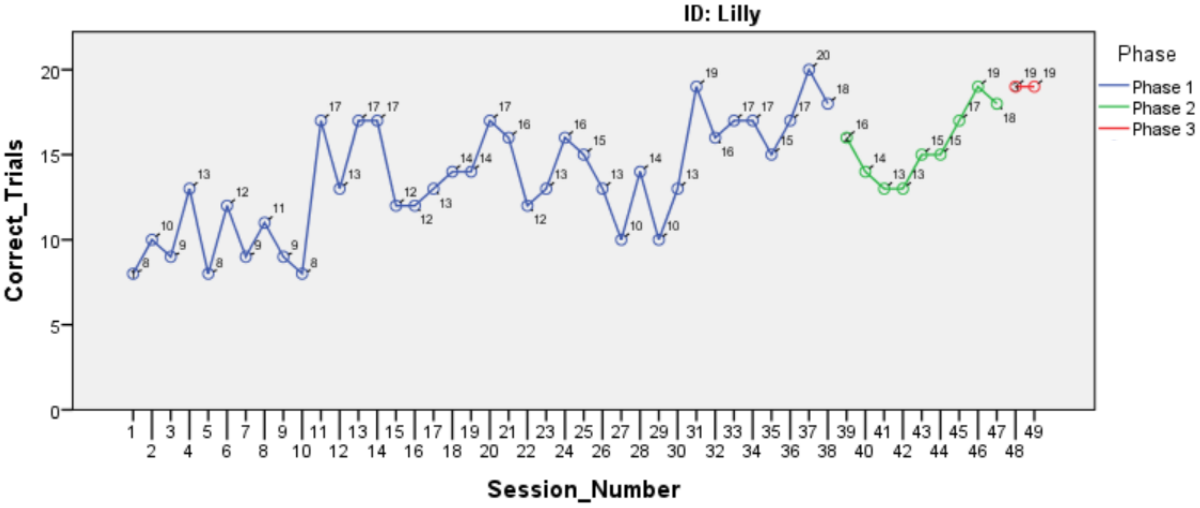

Figure 10: Lilly's results, showing the amount of correct trials per session for each phase

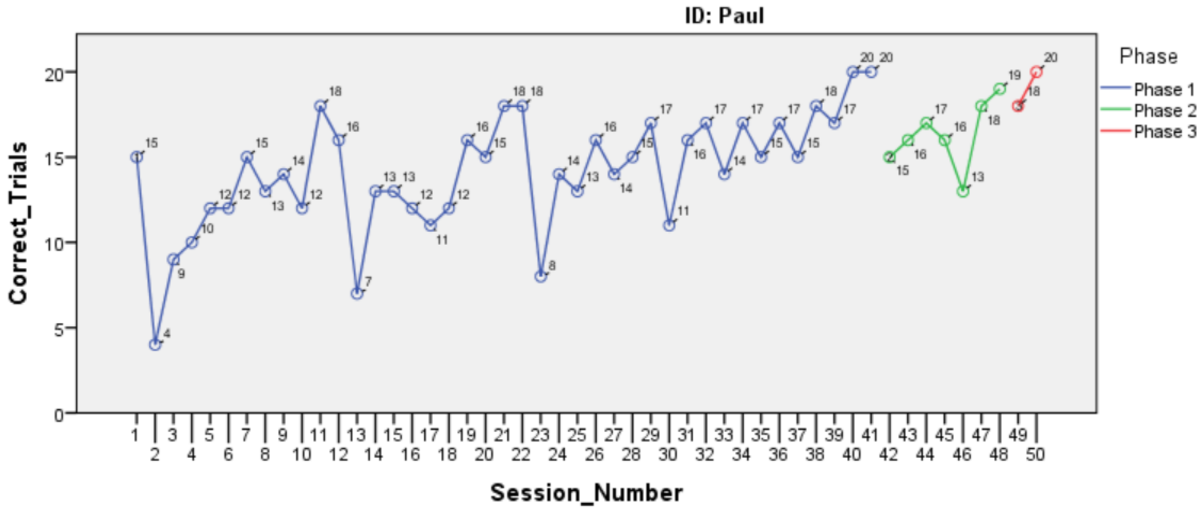

Figure 11: Paul's results, showing the amount of correct trials per session for each phase

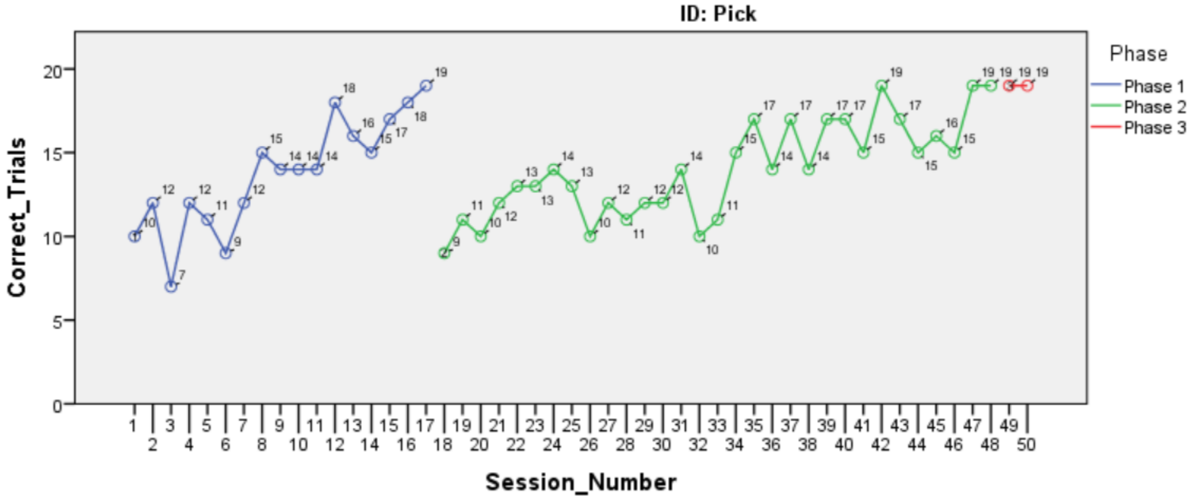

Figure 12: Pick's results, showing the amount of correct trials per session for each phase

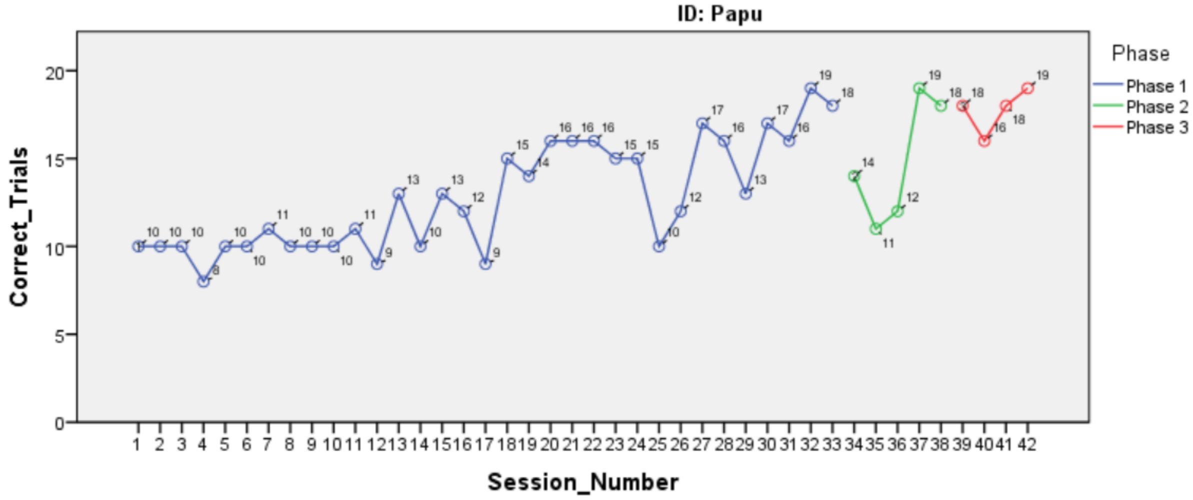

Figure 13: Papu's results, showing the amount of correct trials per session for each phase

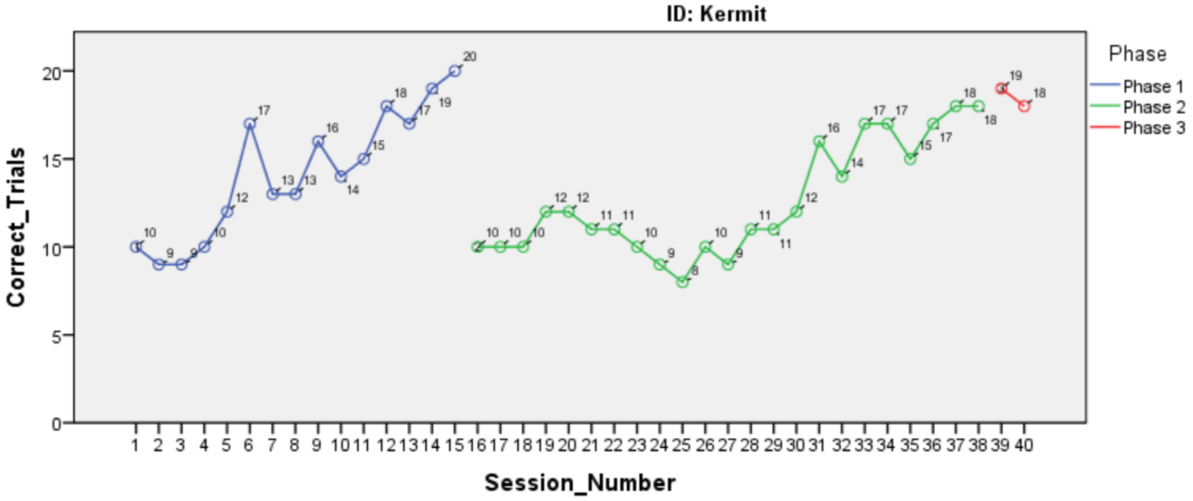

Figure 14: Kermit's results, showing the amount of correct trials per session for each phase

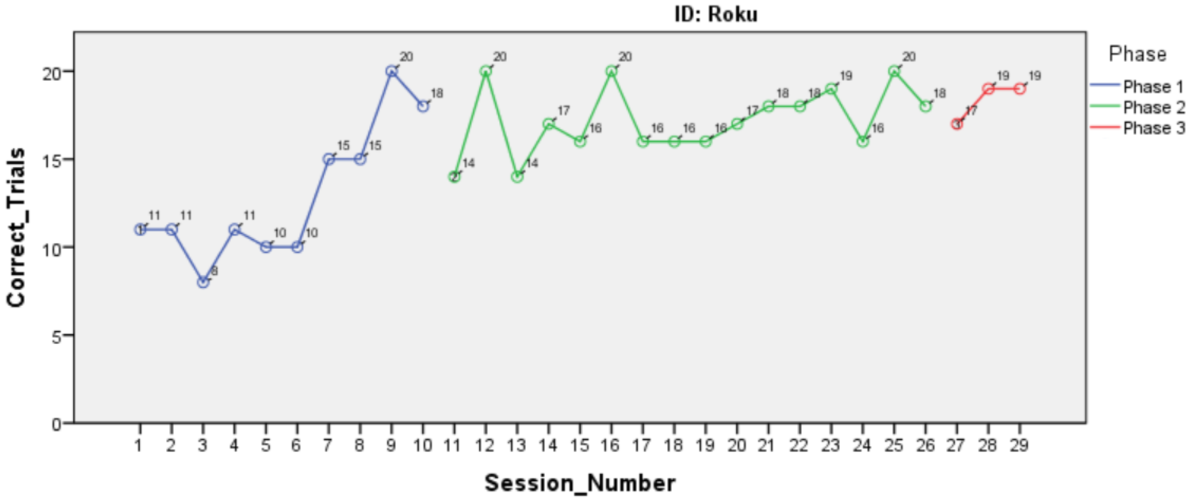

Figure 15: Roku's results, showing the amount of correct trials per session for each phase

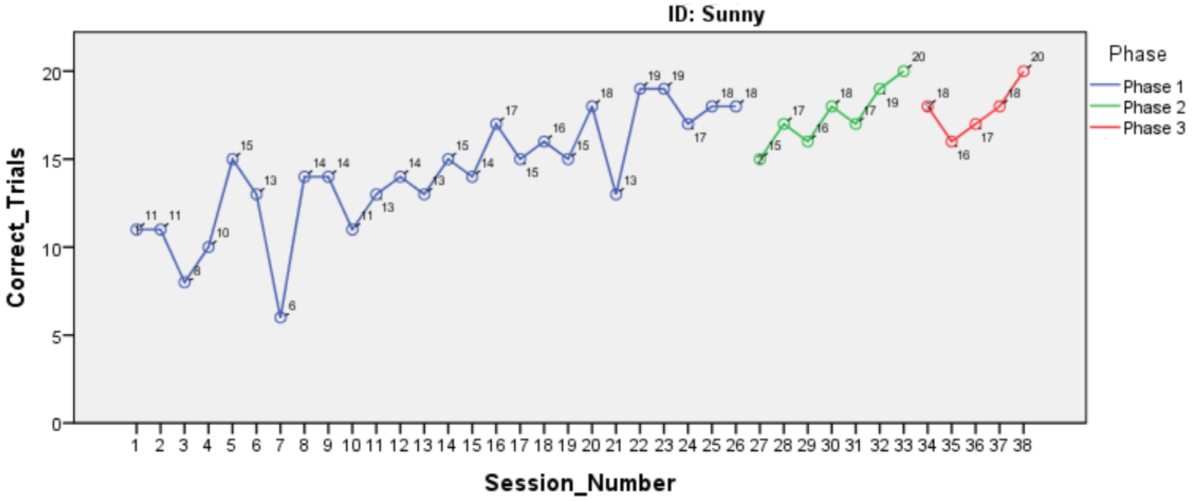

Figure 16: Sunny's results, showing the amount of correct trials per session for each phase

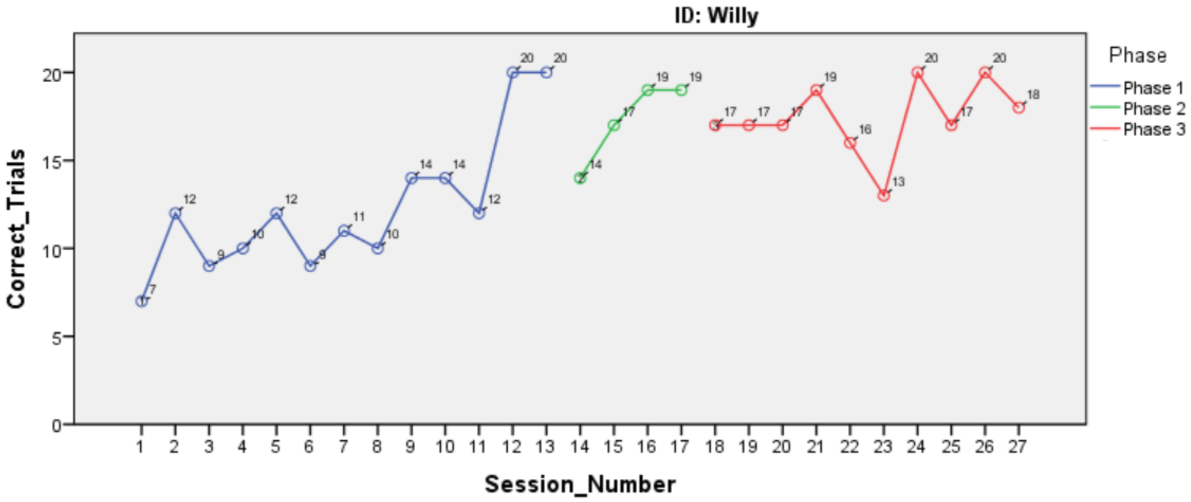

Figure 17: Willy's results, showing the amount of correct trials per session for each phase