Review of Studies on Visual Perception in Grey Parrots (*Psittacus erithacus*): The Müller-Lyer Illusion, Amodal and Modal Completion

Irene M. Pepperberg*

Harvard University, Cambridge, MA, USA

*Corresponding author (Email: impepper@wjh.harvard.edu)


Abstract - Few avian studies on optical illusions are directly comparable to those with humans. Grey parrots that have some referential use of English speech, however, allow for such comparative studies, as these birds can be tested just as are humans, by asking them to describe exactly what they have seen. Here I review two studies, one on the Müller-Lyer illusion (Pepperberg, Vicinay, & Cavanagh, 2008), one on amodal and modal perception (Pepperberg & Nakayama, 2016), that demonstrate similarities between human and Grey parrot perceptual abilities.

Keywords – Grey parrots, Optical illusions, Parrot visual perception, Müller-Lyer, Kanisza, Amodal completion

To better understand the evolution and architecture of visual cognition and visual attention in humans, researchers often study visual processing in other species. Comparative behavioral research—how subjects respond to stimuli—efficiently answers basic questions about how different cognitive architectures process equivalent information: Resultant cross-species comparisons permit identification of both critical, core elements that are preserved across divergent cognitive architectures and those that may be unique to humans (e.g., Feng, Chouinard, Howell, & Bennett, 2016; Jackson & Cormack, 2007). We expect that similar evolutionary demands—visual environments and survival needs—have led to analogous, if not necessarily homologous, solutions concerning certain forms of visual processing. Birds, however, to a greater extent than humans, must navigate in three dimensions, identify and respond quickly to obstacles and predators—demands that match or exceed those placed on primate visual systems—such that a great deal of processing resources must be allocated to vision. How does the avian brain—one that is physically smaller and somewhat differently organized from, but that is still evolutionarily similar to, that of primates (e.g., Jarvis et al., 2005)—process illusory information and what might such information tell us about human visual perception?

Before our work (Pepperberg & Nakayama, 2016; Pepperberg, Vicinay, & Cavanagh, 2008), avian studies on optical illusions were not directly comparable to those on humans. For example, most research that examined optical illusions seen by chicks (Winslow, 1933), pigeons (e.g., Aust & Huber, 2006; Fujita, Blough, & Blough, 1993; Lazareva, Wasserman, & Biederman, 2007; Nakamura, Fujita, Ushitani, & Miyata, 2005, 2006), doves (Warden & Baar, 1929), and both starlings and finches (Dücker, 1966) generally involved training birds to distinguish a particular stimulus and then tested their recognition of similar patterns. Statistical significance in these studies often depended on, for example, averaging over hundreds of trials of pecking/touching behavior to a very limited set of choices and thus results were often highly variable and dependent upon the details of the
experimental design (reviewed in Pepperberg & Nakayama, 2016; Pepperberg et al., 2008). Subjects could not respond in ways similar to those of humans; that is, vocally answer to direct questioning about what exactly had been observed. Although vocal responses are not the only way to examine nonhuman abilities, such responses are optimal. Interspecies communication directly states the precise content of the questions we ask—an animal need not determine the nature of the question through trial-and-error. Interspecies communication is also an open, arbitrary, creative code with enormous signal variety, enabling an animal to respond in novel, possibly innovative ways that demonstrate greater competence than required responses of operant paradigms and allows researchers to examine the nature and extent of information an animal perceives. The technique also allows rigorous testing that avoids expectation cueing: Subjects can be made to choose responses from their entire repertoire rather than from a subset (usually two choices) relevant only to a particular topic. Thus, the study of optical illusions in Grey parrots, a species that has limited use of referential speech (Pepperberg, 1999) and as a consequence can respond vocally, was of great interest on its own in addition to its ability to participate in visual studies.

Grey parrots make particularly useful subjects for such studies of a number of reasons. Despite the phylogenetic distance of birds and primates, strong parallels in both cognitive and perceptual capacities have been documented for Grey parrots and humans. Like primates, they are K-selected, and compared to most birds other than corvids, are relatively larger-brained (Olkowicz et al., 2016), longer-lived subjects. Grey parrot cognitive capacities have been shown to be equivalent to (and in one case, beyond) those of the great apes and marine mammals (e.g., Pepperberg, 2012; Pepperberg & Carey, 2012). Parrots’ perceptual traits are likely to have been selected by evolutionary pressures somewhat similar to those of early primates (see Demery, Chappell, & Martin, 2011): Field studies (May, 2004) reveal that Grey parrots forage over large distances (60 km/day) in multiple habitats (e.g., juicing bark and eating fruit in forest trees, consuming herbs and soil at ground-level savannas); need to be alert for aerial and terrestrial predators; and split their time among large flocks, small groups, and isolated pairs. All these activities involve visual processing and a great need for visual accuracy. Furthermore, as noted above, these birds can acquire referential use of labels for objects, colors, shapes, number, and even absence of attributes (Pepperberg, 1999, 2012), which allows them to respond just as humans to questions about what they see in the world.

Given that the goal of the visual system is to process information accurately—often for survival—my colleagues and I thus did not expect qualitative functional differences across species—avian or human—that similarly depend on vision, but significant differences do exist in the organization of visual systems across species. Therefore, one of the goals of our studies on Grey parrot responses to visually-dependent tasks was to test if differently-organized brains do indeed process certain types of information similarly or if differences, if any, appear at various levels of processing. Specifically, in humans, the lemnothalamic pathway is highly developed with respect to vision; for at least one avian species, the pigeon (Columba livia), the collothalamic visual pathway is the more highly developed (see Nguyen et al., 2004; Shimizu & Watanabe, 2012). Pigeons’ visual system, including right versus left eye dominance on some tasks (Yamazaki, Aust, Huber, Hausmann, & Güntürkün, 2007), may explain why they favor a featural strategy in visual scene analyses and sometimes fail to recognize optical illusions (e.g., Prior & Güntürkün, 1999; Sekuler, Lee, & Shettleworth, 1996; see excellent review of pigeon vision in Regolin, Marconato, & Vallortigara, 2004). Chicks, however, use the lemnothalamic visual system for tasks involving frontal viewings (Deng & Rogers, 1997, 1998a, b), more like humans, and show a more balanced and complementary use of both eyes, which may assist in their more global systems of visual analysis (Regolin et al., 2004, but also see Tommasi & Vallortigara, 2001, who suggest that input from different eyes allow for use of multiple strategies to solve the same task). Although the visual range of only a single parrot species, a Senegal (Poicephalus senegalus), has been studied (Demery et al., 2011), the parrot system likely resembles that of the chick more than the pigeon (G. Harrison, pers. comm., 2006). Psittacine UV color perception, however, is likely more closely related to that of the pigeon than to the chick (Carvalho, Knott, Berg, Bennett, & Hunt, 2011).

With this information in mind, my colleagues and I decided to examine how Grey parrots, who could respond vocally as do humans, would process a variety of optical illusions. Do parrots “fall” for the same optical illusions as do humans? Do they understand depth cues likely needed to perform amodal completion? Studies on such topics would provide behavioral data for direct comparison with humans and a basis for future studies on more complex types of visual processing. Because of the procedural differences, only indirect comparisons would be possible with other avian species (discussions of such comparisons are in Pepperberg & Nakayama, 2016;
Müller-Lyer Illusion (summarized from Pepperberg et al., 2008)

The Müller-Lyer illusion was chosen as the initial task because it is well-represented in the scientific literature, in both human and nonhuman studies. Müller-Lyer (1889, 1981) first described how humans either underestimated or overestimated the length of a line that had arrows attached, respectively, either inwardly < > or outwardly > <. Many researchers have proposed theories to suggest why humans see the illusion as they do (reviewed in Feng et al., 2016; Pepperberg et al., 2008), but even now, no definitive explanation exists. Studies on most nonhumans, which might suggest the roles of environment and/or evolution on susceptibility, are however likely to be affected by the intensive training procedures required in order to obtain data from the candidate species (see discussion above; also Rosa Salva, Sovrano, & Vallortigara, 2014, for a discussion on the importance of avoiding intensive training). Having a subject that, like humans, would have a vocal repertoire from which it could choose any utterance (from over 100 possibilities), and could simply be asked to describe what it sees, on only a few trials per type of stimulus, without any prior training on any materials related in any way to those stimuli, would avoid these issues. Even subjects that are tested on spontaneous preferences procedures usually have, unlike our parrot, only very limited choices (usually between two alternatives; e.g., Wood & Wood, 2015) and usually at least 100 subjects must be tested on the same stimuli in order to achieve statistical significance (e.g., Lea, Slater, & Ryan, 1996). The point of this review is to emphasize the value of interspecies communication for examining comparative visual perception.

Subject

Alex, a male Grey parrot (Psittacus erithacus) who was 30 years old by the end of the study, had been the subject of cognitive and communicative studies for 29 years (e.g., Pepperberg, 1999, 2012). Testing locations and living conditions when neither testing nor training were in progress are described in Pepperberg and Wilkes (2004). Food and water were available at his vocal request at all times. In this study he used previously documented abilities to produce English color labels [Pepperberg (1999): rose (red), green, purple, blue, yellow, orange, grey (charcoal to black)], and respond to questions concerning the concept of relative size and absence of size differences (Pepperberg & Brezinsky, 1991). Training techniques used to establish his communication skills have been described in detail previously (e.g., Pepperberg, 1981, 1999).

Initial Procedures

All of Alex’s prior experience on size difference exclusively involved 3D objects (e.g., woolen pompons, wooden blocks, plastic keys), a task on which he was very successful (Pepperberg & Brezinsky, 1991), and all stimuli would now involve 2D printed figures entirely unrelated to the original training and testing materials. In fact, his ability even to label 2D representations of 3D objects was still unknown. Thus, he was first shown sets of two differently colored 2D horizontal lines and asked to identify the bigger or smaller one or to state “none” if they were equally sized (for details, see Pepperberg et al., 2008). At the time, very little information existed on psittacine binocular overlap or visual acuity, so we chose to work monocularly. Such familiarization was necessary to direct his focus to the 2D stimuli. We kept these trials to a minimum (again, for details, see Pepperberg et al., 2008). We were not concerned with his accuracy as we did not want to train him on this task by rewarding and thus reinforcing particular responses to these queries; we wanted to avoid cuing him about tests to follow. We simply tracked whether he attended to the stimuli. Alex was also given a visual acuity test (see Pepperberg et al., 2008), thus ensuring that his vision would be adequate for the task we intended to use.
Test Stimuli

We tested Alex on the Brentano variation of the Müller-Lyer illusion (Figure 1). This version, equivalent to the presentation of two separate figures (e.g., Sadza & de Weert, 1984), ensured he focused on both illusions simultaneously. By using two central line colors and differently colored arrows, experimenters could query him and receive his answers vocally. (NB: For humans, color can reduce illusion strength with ~7–10% reduction; see Miyahara, 2006; Mukerji, 1957; Sadza & de Weert, 1984.) Use of two different central line colors also allowed repeating the query (“What color bigger/smaller?”) while requiring several different answers (one of seven color labels or “none”), which decreases inadvertent cuing or training to a specific response (see below). Requiring responses with respect to both bigger and smaller also forced Alex to attend to and interpret each question individually, unlike most other nonhuman subjects.

Figure 1. Alex and a stimulus. Before a trial began the stimulus would be positioned for monocular viewing.

Initial Brentano figures (e.g., Figure 2) were constructed so as to address several concerns (additional details in Pepperberg et al., 2008). Sizes of arrows (1 mm x 2.5 cm) and central lines (10 cm x 4 mm) needed to be clearly seen against the background. Central lines, half of one color and half of another, were drawn to appear in front of the arrows to maximize the illusion (Miyahara, 2006). Black arrows, pitched 45º from the horizontal, were based on optimality found in previous studies (e.g., Sekuler & Erlebacher, 1971). To avoid bias, half the stimuli were as in Figure 2 and half with arrow placement reversed. Because parrots’ abilities to see in the UV may cause differences from humans in their color perception (Bowmaker, Heath, Das, & Hunt 1994; Bowmaker, Heath, Wilkie, Das, & Hunt, 1996; Carvalho, Knott, Berg, Bennett, & Hunt, 2011; Cuthill et al., 2000; Goldsmith & Butler, 2005; Wilkie et al., 1998), Alex’s labels for specific color samples was examined prior to introduction of test stimuli. Colors were mixed and matched, such that the purportedly correct response (i.e., seeing the illusion) was divided equally over the six colors he could label. Specific color combinations that could be confusing because of his UV perception (Pepperberg, 1999) were eliminated (orange with red or yellow, purple with red or blue).

The stimuli in Figure 2 were used in initial test sessions (see below); subsequent sessions involved other, related sets of figures, all presented on black backgrounds with white arrows (see Figure 1). One such set used the measures described for Figure 2. For a control set, vertical bars replaced arrows to learn if, when the illusion definitely did not exist for human observers, Alex would also fail to see it and respond “none” appropriately. Other sets tested effects of central line (shaft) thickness and angle slant (see Pepperberg et al., 2008, Figure 3, for examples): would Alex’s ability to report the illusion change like humans’ as central lines were emphasized or de-emphasized and arrows approached or receded from perpendicularity (e.g., Coren & Porac, 1983; Sekuler & Erlebacher, 1971)?
Test Procedure

Test questions were presented intermittently either during free periods or sessions on unrelated topics (e.g., number concepts). Thus, although the stimuli clearly alerted Alex to the nature of the task, he could neither focus on nor expect that all questions in a session would be of one type and thus had to attend to the specific query being posed. Consequently, his responses had to be chosen from his entire repertoire and from among numerous possible topics during each session. Such a procedure also eliminated expectation cuing on the part of the experimenter: An examiner who, for example, poses a series of similar questions may come to expect a particular answer and unconsciously accept an indistinct (and by our criteria incorrect) response (Pepperberg, 1999).

Details of test procedures, including details of other precautions against inadvertent cuing, including use of a blind coder, have been described comprehensively elsewhere (e.g., Pepperberg, 1981, 1992, 1994; Pepperberg et al., 2008). I summarize only a few of the many controls here. Because Alex had not been trained on this task, no overlap occurred between training and testing situations; moreover, because color label training occurred years before, by students no longer present, he could not have used trainer-induced cues to specify a given label. The experimenter could not cue Alex as to which side (the ‘bigger’ or ‘smaller’) was correct in a given trial by eye movement because Grey parrots do not respond to human gaze at the distances involved (Giret, Miklósi, Kreutzer, & Bovet, 2009; Pepperberg, Koepke, Livingston, Girard, & Hartsfield, 2013). Because stimulus direction varied and queries varied as to bigger or smaller (i.e., questions were neither massed nor necessarily alternated, but could be either the same or different on a given day, and often only one was asked on a given day), Alex could not learn to respond, for example, only to colors associated with arrows pointing in a particular direction, or to the right or left side. Also, because his specific response (one of six colors or “none”) varied from trial to trial, he could not be trained or cued as to one proper response, and experimenters praised him equally no matter what he said (unless he was being noncompliant; Pepperberg & Carey, 2012; Pepperberg & Gordon, 2005). Moreover, he was unlikely to have been trained or been cued as to the association between arrow direction and size, because his first actual error (giving a response in the incorrect direction) did not occur until trial 22.

Testing occurred in two sets of 32 trials each, with a month break between sets; only 2 to 10 trials were given per week. Breaks occurred for experimenter absences and scheduled dishabitation periods. The latter, for 1 to 3 weeks at a time, examined whether what appeared to be habituation to the illusion (or simple boredom) was indeed occurring, and if a break in testing for a defined period would reinstate a response. These breaks occurred after trials 16, 18, 21, 24, 29, 43, 47, 53, and 60. During the month break between sets, one trial with perpendicular arrows served as an additional control. A trial was generally repeated in a session only if Alex’s initial answer indicated noncompliance (e.g., a color not being presented, Pepperberg & Gordon, 2005) or an indistinct response (e.g., Pepperberg, 1981, 1987). Thus, the number of times stimuli were presented depended on Alex’s behavior. If his response was noncompliant or indistinct, the examiner removed the tray, turned his/her head, and said “No!” Presentation continued until an appropriate identification was made or four attempts occurred; all responses were recorded (see Pepperberg & Carey, 2012; Pepperberg & Gordon, 2005). If he produced either color on the stimulus, he received praise and could request a reward. When, approximately halfway through the first set of test trials, Alex
began to respond “none” to every stimulus, for several trials (potential habituation/boredom) he was asked, “Are you sure? What color bigger/smaller?” to ensure he was not simply failing to attend; on such trials he was queried up to four times; only his first response was recorded. As in previous studies (e.g., Pepperberg, 2006), a small percentage of sessions was videotaped and subsequently analyzed to ensure interobserver reliability.

**Results**

Alex’s data were scored as ‘illusion reported’ if he named the shaft color that human observers would report, as ‘no illusion’ if he reported “none,” and ‘opposite the illusion’ if he reported the color opposite to the illusion response. He reported the illusion in approximately 65% of the trials where human observers would have reported the classic illusion and he showed a lessened or absent illusion in control trials where humans would not have reported the illusion (see Figures 4 and 5 in Pepperberg et al., 2008). If we discount likely habituation, however, his scores are much higher. For example, he reported the illusion on 8 of his first 9 experimental trials in which he would have been expected to do so, and consistently reported the illusion after a break in testing. He failed to report the illusion on only 18/50 trials where he would have been expected to do so, and of those 18 trials, 12 could be attributed to habituation (or boredom), leaving only a 12% failure rate. Such a rate is consistent with his overall accuracy in color labeling (80 – 85%; Pepperberg, 1999). Notably, he responded as expected on all control trials, failing to report the illusion when perpendicular bars replaced the arrows, and reporting “grey” (his label for black) when asked “What color smaller?” when stimuli were as in Figure 2 and the arrows were indeed the smallest lines. These latter trials ensured that he was fully attending to the entire stimulus.

Shaft width indeed affected Alex’s responses. As shafts became more prominent and (at least for humans) directed attention away from the arrows, he was less likely to report the illusion and more likely to habituate. He showed a consistent illusion response with the thinnest shafts (2 mm, 8/11 illusion responses: Of the three failures to report the illusion, only one was a reversal and two involved likely habituation). Responses were again in the illusion direction with the intermediate shaft width and thin arrows (4 mm shaft, 1 mm arrows, 9/14; of the five failures, four likely involved habituation) and with the intermediate shaft and equal width arrows (both 4 mm, 12/20; of 8 failures, 6 again may have involved habituation). For these stimuli, breaks in testing usually restored his response to the illusion. However, he never reported the illusion for 6 mm shafts and 1 mm arrows (0/5) even after breaks from testing of 2 – 3 weeks. Alex did not appear to be responding to the overall lengths of the stimuli rather than to the illusion, as his responses varied from 0% to 73% for stimuli with exactly the same lengths but varying ratios of shaft/arrow widths.

Pitch of the arrows from the horizontal affected responses. Alex reported the illusion more often when angles were more acute and less likely or not at all for less acute angles. On all trials when arrows were replaced with perpendicular bars, Alex responded “none” as do human observers. When the arrow angle approached the perpendicular (i.e., at 80º from the horizontal), he reported the illusion on only 1 of 5 trials. At the other extreme, when the arrows were at 10º, Alex reported the illusion on 3 of 5 trials even though the shaft width was 6 mm, a width for which he did not report an illusion when arrows were at 45º angles.

With respect to interobserver reliability, four videotaped trials were digitized and presented auditorially only to four naïve adult humans who had never interacted with Alex. They were simply asked to repeat what they thought they heard the parrot say. All four agreed on all four samples, and their responses agreed with what the experimenter had recorded (i.e., 100% reliability).

**Discussion**

Alex did, under a variety of conditions, perceive the Müller-Lyer illusion in ways similar to human subjects. Clearly, neither human neuroanatomy and processing, nor human binocular vision, are required for such perception. However, reasons for his performance, as for humans, are not entirely clear. Attention—for example, to global versus local cues—likely plays some role as might visual experience. Given that human studies suggest the involvement of higher visual areas in processing of the illusion (e.g., Weidner, Boers, Mathiak, Dammers, & Fink, 2010), the data may have implications for further comparative research (see Vetter & Newen, 2014).

Attentional issues deserve considerable attention themselves. Some explanations for the illusion claim its origin lies in the arrow-out version being overall larger than the arrows-in version and/or argue that the mechanism
is likely attention to global versus local cues (e.g., see Day, 1989; Mundy, 2016; note Pressey, 1967); our data suggest these issues are somewhat separable. Alex’s responses varied not with overall length of the stimuli (arrows plus shaft) but rather with the ratio of shaft/arrow width as well as shaft/arrow angle, suggesting how emphasis on certain local cues can affect responses. Had he responded solely to the total length of the stimuli (i.e., learned to respond to “What color bigger?” for outward arrows and “What color smaller?” for inward arrows), he would not have failed to see the illusion in stimuli where the shafts were six times wider than the 45º arrows and would have seen the illusion more often than he did for stimuli in which the arrows and shafts were equally wide (4mm for both, 45º arrows). These findings are, however, consistent with attention being focused on certain local cues, as the 6-mm shafts tend to draw attention strongly away from the effects of the arrows, at least until the arrows themselves were extreme (e.g., 10º) and acted to reinstate a response to the illusion—although less strongly than in the original form tested (4-mm shafts, 1 mm and 45º arrows). Humans may respond similarly when shafts are emphasized over arrows (Coren & Porac, 1983; Gardner & Long, 1961). Habituation also obviously involves overall attention: Several researchers, from Mountjoy (1958) and Rudel and Teuber (1963) through Predebon (2006) have claimed that repeated exposure decreases human reporting of the Müller-Lyer illusion (reviewed in Miller & Al-Attar, 2002); such an effect could have also occurred for Alex—notably, his responses decreased after a number of successive trials and were often reinstated after a break from testing.

Monocular presentation may also decrease human perception of the illusion (Mountjoy, 1960; but note Gregory, 1963, 1968). Because Alex was presented with the stimuli so as to ensure monocular viewing, his response strength may have been affected. Given that at the time little was known about parrot binocular overlap—and nothing is still known specifically for Grey parrots—monocular presentation seemed optimal.

Human experience with what is considered a “carpentered world”—one with many examples of right-angled, perpendicular-parallel intersection—appears to engender stronger responses to the illusion (Ahluwalia, 1978; Segall, Campell, & Herskovitz, 1966). Supposedly, the illusion is not innate but develops (or degrades, see Brosvic, Dihoff, & Fama, 2002) during environmental interaction. The issue appears to involve seeing the 2D illusion as representing a 3D impression of, for example, walls and corners. (NB: a subsequent study described below suggests that Grey parrots are capable of 2D/3D transfer, but whether that occurred in our Müller-Lyer experiment is unclear. For comments with respect to the validity of this argument, see Deregowski, 2013.) Alex, however, was hatched and raised in captivity, with exposure to the same visual experiences as a human child. Thus, studies on wild, or at least wild-caught, parrots deprived of such experience would be required to determine the effects of environment on a nonhuman visual and cortical system.

Given that no neurobiological studies have been performed on Grey parrots with respect to their viewing the illusion, anything stated about their neural processing mechanism is purely speculative (see Pepperberg et al., 2008). Studies on humans using MRI and magneto-encephalography (MEG) suggest involvement of the lateral occipital cortex (LOC), the right superior and inferior parietal cortex, the right inferior frontal cortex and the right intraparietal sulcus (IPS), among other areas (Weidner & Fink, 2008; Weidner et al., 2010). Weidner et al. (2010) suggest that

“…lateral occipital and inferior temporal regions together with dorsal stream areas play an essential role in the generation of the Müller-Lyer illusion. This is consistent with theories considering a relevant contribution of higher visual areas to the generation of the Müller-Lyer illusion.” (p. 1594)

Whether the human LOC relates to the lemnothalamic or collothalamic pathways or a combination of both in parrots is not known, but birds’ visual brain areas function in a manner corresponding to the LOC (Butler, Manger, Lindahl, & Arhem, 2005). The avian nidopallium caudolaterale (NCL) and the dorsolateral corticoid (CDL) areas, which respectively are likely homologues of the posterior parietal cortex and parts of the prefrontal cortex (Butler et al., 2005; see also Jarvis et al., 2005, 2013), would probably be involved in such processing. The suggestion is that despite what are likely differences in the avian-human visual systems (i.e., processing of visual input), some similar level of higher cognition is involved in processing the illusion in both humans and Grey parrots and that these birds, particularly because they can respond vocally as do humans, may be good models for understanding the cognitive processes underlying such illusions.
Subjective contours involve ecologically relevant stimuli. Humans often fill in missing parts to facilitate the perception of objects in their environment. Early Gestalt psychologists (e.g., Kanizsa 1979, 1987) described two of the most common forms of this behavior: amodal completion, when the object of interest is occluded by some other item (Figure 3A) and modal completion, when the object is actually illusory but nevertheless appears to exist (Figure 3B). Nakayama, He, and Shimojo (1995), using psychophysical tests, showed that such processes are likely localized within the visual system insofar as purely visual tasks such as motion perception, texture segregation, attentional allocation, and object recognition depend critically on these processes. Furthermore, neurophysiological recordings in nonhumans (e.g., Bakin, Nakayama, & Gilbert, 2000; von der Heydt, Peterhans, & Baumgartner, 1984; Zhou, Friedman, & von der Heydt, 2000), demonstrated that neurons in the visual cortex show responses indicating the existence of such processes. Because the problems posed by occlusion and lack of luminance transitions across object boundaries are possibly universal—many other creatures must experience this problem in their daily lives (e.g., Lea et al., 1996; Regolin & Vallortigara, 1995; note Tvardiková & Fuchs, 2010) - ‘filling in the blanks’ (perceptual completion) is likely a fundamental visual process. For example, processing partial clues about a potential predator and reacting is safer than not, even if some false alarms incur costs.

Just as noted for the Müller-Lyer illusion, however, showing that nonhumans’ responses to such visual stimuli can be interpreted unambiguously has not been easy. Many studies argue that various nonhuman subjects respond appropriately to 2D objects similar to those in Figure 3 (e.g., DiPietro, Wasserman, & Young, 2002; Horridge, Zhang, & O’Carroll, 1992; Kanizsa, Renzi, Conte, Compostela, & Guerani, 1993; Martin-Malivel, 2011; Regolin & Vallortigara, 1995; Van Hateren, Srinivasan, & Wait, 1990), but all of these studies are subject to a variety of alternative interpretations: Responses may have been based on, for example, local cues, mass/number, or stimulus generalization (see Pepperberg & Nakayama, 2016). Importantly, subjects in many, but not all of these studies (for an exception, see Regolin & Vallortigara, 1995) are not directly questioned about what they see, but undergo significant training prior to testing, engendering the same issues raised above with respect to the Müller-Lyer illusion. In contrast, another Grey parrot, Griffin, who, like Alex before him, knew labels for various colors and shapes based on 3D objects, could simply be asked appropriate questions about such figures, just as would be done for humans. He had not, however, had experience with 2D stimuli, or with any non-regular polygons, so we would be testing his capacities to transfer his knowledge to identifying such items as well as examining his abilities with respect to modal and amodal completion.

Griffin, a male Grey parrot (Psittacus erithacus), 16 years old at the beginning of the study, had been a research subject since he was 7.5 weeks old (e.g., Pepperberg, 2012; Pepperberg & Hartsfield, 2014; Pepperberg et al., 2013; Pepperberg, Willner, & Gravitz, 1997). His living conditions matched those of Alex (Pepperberg & Wilkes, 2004). We used his abilities to produce and comprehend labels for various colors [red (rose), blue, green, yellow, orange, purple] and shapes (regular 3D polygons labeled “1-”, “2-”, “3-”, “4-”, “6-”, “8-corner”). He could also identify, with standard English number labels, sets of clicks, thus demonstrating some quantification skills. He
received basically the same type of training to establish communicative competence as did Alex. And, like Alex, he also grew up in an environment more similar to that of a preschool child than a laboratory subject, with the ability to identify different objects, foods, and locations.

**Initial Procedures**

Like Alex, Griffin had not been trained on 2D stimuli; here we were particularly interested whether he could transfer directly from training on 3D objects to testing on their 2D representations. Such transfer was a prerequisite for success on amodal and modal completion. To examine Griffin’s transfer abilities, we needed a 2D stimulus that would not interfere with the amodal and modal completion tasks, but would test if he could focus on the appropriate target. We chose a square that occluded a disk he could not identify (Figure 4; he could neither label circles nor black).

We introduced these stimuli (for all colors he could label) during a period of noncompliance, similar to those experienced with Alex (e.g., Pepperberg & Gordon, 2005), where he was using a single label to respond to every question, whatever the topic. For Griffin, novelty can often break this behavior pattern. He was thus queried, “What shape green/blue/yellow/orange/rose?” until, on the thirteenth trial, he responded appropriately.

![Figure 4](image)

*Figure 4. A ‘dummy’ occlusion stimulus. Green object was the target of the question (“What shape green?”). Colors and shapes varied over trials.*

**2D Test Stimuli**

To test for amodal completion, we asked Griffin to label colored regular polygons he could identify occluded by black disks he could not label (e.g., we asked “What shape rose?” as in Figure 5A; details in Pepperberg & Nakayama, 2016). We also used other known black shapes as occluders (see, e.g., Figure 5l, Pepperberg & Nakayama, 2016), to determine if being able to label occluders would distract him from the task at hand. As controls, we also tested non-occluded figures that contained exactly the same parts but did not allow for amodal completion (e.g., Figure 5B); we called these figures detached probes. If Griffin understood amodal completion, he would label Figure 5A as “3-corner”, but Figure 5B as “4-corner”. An oval or circular shape was cut around every stimulus so the shape of the paper (“4-corner”) would not influence Griffin’s response. Colors were those used with Alex to avoid issues involving Grey parrot color perception. We used 38 figures (33 for amodal completion, 5 detached probes). The complete set is in Pepperberg and Nakayama (2016; see Supplementary Material A in that paper).
Figure 5. (A) Amodal stimulus; (B) Detached probe.

For modal completion—also known as subjective contours or Kanizsa figures—black ‘pac-men’ formed regular polygons on printed colored paper (Figure 6A); colors chosen as before. Controls (“probes”) involved placing additional circles or ‘pac-men’ near the Kanizsa figure so Griffin could not simply quantify black objects (e.g., Figure 6B). Here a random amorphous non-polygonal shape was cut out around the target stimuli so that the shape of the paper would again not factor into his response. We presented 38 stimuli (27 regular Kanizsa figures; 9 probes). The full stimuli set is in Pepperberg and Nakayama (2016; Supplementary Material A in that paper).

Figure 6. (A) Modal stimulus; (B) Modal probe.

Test Procedure

Testing generally followed the same procedures used for Alex. Griffin was shown stimuli monocularly and asked to report vocally what he was seeing. As for Alex, numerous controls of the types described above were in place to avoid various types of cuing (see Pepperberg et al., 2008; Pepperberg & Nakayama, 2016). And as before,
several trials were videotaped for subsequent interobserver reliability. For additional details, see Pepperberg and Nakayama (2016).

Unlike nonhuman subjects in other studies, who undergo considerable training and whose training trials are generally repeated during testing (e.g., DiPietro et al., 2002; Nagasaka, Brooks, & Wasserman, 2010), all of Griffin’s trials were true test trials, as he had never seen any of these standard test stimuli, probes, or detached probes prior to this study. To ensure that no trials of any type were administered consecutively (which could encourage learning and/or stimulus generalization), stimuli presentations were randomized with respect to modal/amodal task, shape, size, and color, and all such trials were intermixed with trials on other types of tasks (e.g., exclusion; Pepperberg et al., 2013). No specific stimulus (confluence of color/shape/size/arrangement of occluders or pac-men) was ever presented in more than one trial. A shape query could be repeated for these very different arrangements, but never for more than 5–6 times in these differing configurations, and generally far fewer. As for Alex, queries about a particular stimulus could be repeated during the single trial during which it was presented if Griffin engaged in noncompliant behavior. However, because additional exposure to stimuli increases the chance of success (DiPietro et al., 2002; Nagasaka et al., 2010), we intentionally limited the number of overall trials. Griffin received no more than 1–2 trials on any given day; thus he could not engage in win/stay-lose/shift behavior (i.e., perseveration).

One criterion for confirming modal and amodal completion was that Griffin respond appropriately not only to the standard stimuli but also to probe and detached probe trials on their first and only presentation so no training could possibly be involved. This criterion was different from that of most other studies on nonhuman subjects, which repeatedly present identical probes during testing (e.g., Nagasaka & Wasserman, 2008; Nakamura, Watanabe, Betsuyaku, & Fujita, 2011), and either reward subjects for all probe trials (potentially encouraging guessing) or, after decreasing primary rewards to a set percentage similar to the proportion of probe trials, for none of the probe/test trials (potentially discouraging possible correct attempts). In these studies, nonhuman subjects may have their testing stopped and undergo retraining if their accuracy decreases in the familiar (training) trials that continue to be presented during testing (DiPietro et al., 2002). Such additional training likely increases the chance of success. We avoided any of these issues by using only a single presentation of each possible probe (see Pepperberg & Nakayama, 2016). We also used a limited number of standard stimuli (also to avoid training). Overall, the percentage of such probe trials in our study was similar to that of other experiments on nonhumans (e.g., 11–25% of test trials within training stimuli; Fujita & Giersch, 2005; Nagasaka & Wasserman, 2008).

Results

Griffin successfully transferred his ability to identify trained 3D stimuli to appropriately label 2D test figures in completion tasks for both amodal (70%; 23/33) and modal (76%; 29/38) trials; he was also 100% correct on all five detached probe trials lacking occlusion. Taking chance at 20%, his performance was statistically significant (binomial test, \( p < 0.01 \), for all types of trials, including detached probes). Even using chance of 50% on the amodal task (one might argue that his response should either be the label of the occluded shape or the number of visible corners of the actual object), his data are still statistically significant (\( p < 0.01 \)). Notably, transfer occurred immediately, with no learning curve: He was correct on his first trials for both amodal and modal completion, and no difference in accuracy existed when comparing scores on his first versus last 15 trials for both types of completion (Fisher exact tests, \( p > 0.2 \)).

Details of his responses can be seen in graphs and figures in Pepperberg and Nakayama (2016); I review only highlights here. Interestingly, Griffin made no errors on squares and hexagons in either completion task, and few on regular 2-cornered polygons; the majority of errors were on 1- and 3-cornered polygons and triangles, which he seemed to confound in both modal and amodal tasks (see discussion in Pepperberg & Nakayama, 2016). In the majority of cases, when was asked a second time about figures on which he erred, he provided a correct response without any additional input from the examiner. However, he was never scored “correct” on such trials.

In amodal completion trials, he responded to the true shape of the figures, even though parts were hidden; he did not simply count corners although his labels for shapes (“#-corner”) would have encouraged such behavior. In contrast, for detached probes, where no occlusion occurred, he labeled the irregular polygons—figures he had never before seen—with the correct number of visible corners. He clearly understood the difference between these two
cases. He very occasionally responded with the shape of the occluder when it was known (counted as an error), but corrected himself when asked a second time.

In modal completion trials, Griffin did not simply quantify ‘pac-men’ or numbers of circles, because in no instance did an error correspond to the quantity of these objects. Nor, for the most part, was he simply responding to, for example, a particular size angle, because he successfully identified a triangle with a 90º angle as “3-corner” (he would have called it “4-corner” otherwise) and also correctly identified a trapezoid as “4-corner” when its size did not encourage attention to local cues (see Pepperberg & Nakayama, 2016). Even his confounds of 1- and 3-cornered objects did not seem to be a consequence of attending only to the 60º angle, because he only once mis-identified a 2-cornered object, which also contained the same angle.

Discussion

Although exclusively trained on 3D objects and given few opportunities to interact with 2D stimuli, Griffin succeeded in tasks of 2D modal and amodal completion. Without training, he therefore not only interpreted 2D images, but also visually inferred the true shape of an even partially visible object or, for subjective contour stimuli, when nothing of the trained image remained. Most other nonhumans animals either do not show these completion phenomena at all or show some degree of success only after having undergone considerable training or after having considerable experience with very closely related stimuli (for detailed discussion of this point, see Pepperberg & Nakayama, 2016). Thus, although both types of completion have been reported in various other nonhuman species, those results are not fully convincing.

What is in question is how similar the experimental design of these other studies are to those given to humans. In nonhuman studies where training was not an issue, success or failure may have involved mechanisms unrelated to the perceptual ones being examined in Griffin, and performance may have actually relied on mechanisms that do not match those of humans (e.g., luminance, aspect ratio; Minini & Jeffery, 2006). In most nonhuman studies, actual tasks differed considerably among the laboratories (for example, with respect to motion, 2D vs 3D stimuli, CRT vs LCD monitors [i.e., flicker fusion effects] and pre-exposure to stimuli; Nagasaka, Lazareva, & Wasserman, 2007; Nagasaka & Wasserman, 2008; Vallortigara, 2006) and such differences were likely responsible for the considerable variation in results (e.g., Aust & Huber, 2006; DiPietro et al., 2002; Lazareva, Wasserman, & Biederman, 2007; Sekuler et al., 1996; Ushitani & Fujita, 2005; Ushitani, Fujita, & Yamanaka, 2001). Pepperberg and Nakayama (2016) provided detailed contrasts between Griffin’s study and those used with other nonhumans (e.g., stimulus generalization, emphasis on local versus global cues, mass and numerical-based responses) and the arguments can be examined therein. Here, I emphasize only two issues, because they related most strongly to Alex’s data in the Müller-Lyer experiment and comparisons with humans: that of visual experience and symbolic representation (summarized from Pepperberg & Nakayama, 2016).

Griffin’s environment was significantly different from that of most (but not all; see the Nagasaka et al., 2010, study on bonobos) nonhumans tested for modal and amodal completion. Human children need considerable interaction with the 3D world in order to develop these abilities (e.g., Kellman & Spelke, 1983); as noted above, from the time of his entrance into the laboratory at 7.5 weeks, Griffin’s 16 years of daily life had been more like that of a human toddler than an experimental laboratory subject. Because he not only consistently interacted with and viewed 3D items but also, at times, saw the same items in ways that would seem more 2D (e.g., the flat surface of a data sheet that was then handled in a 3D manner by a human), he may have less difficulty than other subjects in transferring from 3D to 2D representations (note Stephan, Wilkinson, & Huber, 2013). Moreover, although occluded 3D objects were avoided during training, Griffin would have seen how our hands occluded food items as they were being chopped, food bowls as they were being cleaned, or books, laptop, clothing that students manipulated in his presence. Such exposures were very different from those of most nonhumans, whose experiences were usually to very limited number of shapes: they were generally trained to identify 2D figures on a 2D background, and testing conditions closely resembled those of training. Given that both amodal and modal completion are likely three dimensional processes, requiring an implicit understanding of the 2D stimuli as representing 3D forms (e.g., recognizing that the square in Figure 3A is in a plane behind that of the occluding circle, or that “triangle” in Figure 3B could be viewed as a figure that is in front of, and occluding, three circles; see Figure 10 in Pepperberg & Nakayama, 2016), seeing only a 2D projection of an object would not likely enable an
animal to understand it as a representation of that 3D object. Interestingly, when given a task involving considerable training on 2D figures and forced, two-choice testing on 2D transfer stimuli, baboons failed on amodal completion (Deruelle, Barbet, Dépy, & Fagot, 2000) but did succeed (although only after several hundred training trials and 60 forced, two-choice testing trials) when given both training and testing stimuli that provided background depth cues indicating that the occluder was indeed in front of the targeted object (Fagot, Barbet, Parron, & Deruelle, 2006; see also Nagasaka et al., 2010, on bonobos). Similarly, an amodal completion experiment with the chimpanzee Ai (Sato, Kanazawa, & Fujita, 1997), involved motion of rods behind occluders, which also provided some level of depth cues (however, other, non-amodal completion explanations of their data are also likely; see Pepperberg & Nakayama, 2016).

Griffin, like Alex before him, understood and could use symbolic representation: For over 16 years, he had been taught, in both comprehension and production modes, that vocal labels represented colors, shapes, materials, foods, locations, categories, and to a limited extent (compared to Alex), that single Arabic numerals could represent sets of multiple objects (e.g., Pepperberg & Carey, 2012). Given this equivalence, he would be more likely than most other nonhuman subjects to understand that a 2D form could represent a 3D object. Notably, bonobos in the amodal completion study of Nagasaka et al. (2010) were also adept at labeling and thus symbolic representation. Use of vocal labels also allowed us simply to ask Griffin to report what he saw, much as would be done with humans, on only a small number of trials, in order to establish statistical significance (e.g., Bulf, Valenza, & Simion, 2009). We did not rely on statistical averaging of hundreds of trials (e.g., Deruelle et al., 2000; DiPietro et al., 2002) involving pecking behavior to a limited set of choices (e.g., for amodal completion, between a whole and very closely-related partial figure), but rather relied on vocal English responses: one of a subset of 6 labels chosen from an entire repertoire of roughly 30 labels. In contrast, the hundreds of test trials in other nonhuman studies provide the possible opportunity for new learning. Note, too, that Griffin was never food or water deprived—the cost of an error was minimal, merely lack of a favorite treat, which was quite likely to be gained in a different, usually sequentially presented task. Thus, his responses were not motivated by hunger or thirst.

What is intriguing is that the cross-species data indicate that the capacity could rely on relatively early-level processing and possibly a single mechanism that is widespread across taxa (Ben-Yosef & Ben-Shahar, 2012; Kalar, Garrigan, Wickens, Hilger, & Kellman, 2010; Nakayama et al., 1995)—that is, that the mechanisms are likely similar for the two forms of completion and that parrots likely have the appropriate brain structures for such completion. For example, recent studies (Wokke, Vandenbroucke, Scholte, & Lamme, 2013) as well as earlier experiments (Murray, Foxe, Javitt, & Foxe, 2004) suggest that LOC is strongly involved in human modal completion, and the existence of an avian analog of LOC and analogous processing of such visual material was discussed above. Thus, Grey parrots have again been shown to be good candidates for additional studies on visual perception and direct comparison with humans (again, note Vetter & Newen, 2014).

**General Discussion**

Our data suggest that Grey parrots who have been trained on symbolic representation and that have experienced environments similar to those of young humans could provide particularly valuable data in the study of comparative visual processing. Despite over 300 million years of evolutionary separation between birds and mammals (e.g., Carroll, 1988; Hedges, Parker, Sibley, & Kumar, 1996), both Alex and Griffin perceived various optical illusions in ways very similar to those of humans. Although no neurophysiological studies have been performed on these birds, knowledge of their behavioral data and research on related avian species (e.g., Jarvis et al., 2005, 2013; Olkowicz et al., 2016) suggests that even if systems for visual input differ from those of mammals, similar processing may occur at somewhat higher levels. Our work should provide impetus for future research on the perceptual abilities of Grey parrots and further comparisons among avian and nonavian subjects.

For example, had Alex not died when he did, he would have been tested on the corridor illusion for comparison with humans (see related study by Yonas, Granrud, Arteberry, & Hanson, 1986) and other nonhuman subjects (e.g., Kelley & Endler, 2012). Notably, Griffin understands concepts of bigger versus smaller, and is currently being trained to respond “none” to the absence of a size differential; our plan is to examine his abilities on the corridor illusion once that training has been completed. Such a study is particularly interesting because
flighted creatures must encode depth. Do birds in particular, with their different visual systems, use the same cues as human and nonhuman primates? Or have evolutionary pressures of a flighted niche emphasized different cues? Comparisons on the corridor illusion will help us understand what aspects of human cognition and attention are specific to human neural architecture and what aspects are driven by environmental factors to provide common solutions in different brains.

In a rare departure from our dependency on vocal responses, a colleague and I (Pepperberg & Pailin, 2017) have also tested Griffin on visual working memory and visual working memory manipulation. The goal was to determine whether abstract mental manipulations based on visual processes were uniquely human. In this study, we compared humans and Griffin on a “shell game”-like task. Multiple (2 to 4) colored objects were presented briefly, then occluded. On static trials, items remained stationary (storage). On dynamic trials, pairs of occluders swapped positions up to four times (manipulation of color-location bindings). Participants indicated the color that was expected under a cued occluder. For both Griffin and humans, performance on static trials were consistent with 3 – 4 storage limits, and diverged somewhat only on the most difficult dynamic trials. Our data do not address the mechanisms that Griffin may have used; that is our next challenge.

Overall, these findings raise other interesting questions. Might the same processes we see in parrots be observed in future studies of nonhuman primates and other avian subjects with similar experiences? Might pigeons that live in aviaries and that experienced constant free flight and extensive interactive experience with many different objects succeed in tasks as well as did Alex and Griffin (note Stephan et al., 2013, for some suggestion of such possibilities). Could the mechanisms used by Grey parrots be mathematically modeled to extend to machine and robotic learning? Whatever the answers to these questions, the data suggest that Grey parrot visual processing should be tested with additional illusions and related tasks.

Acknowledgments

Research reviewed in this manuscript was supported in part by a National Science Foundation grant BCS-0920878 (Nakayama, PI; Pepperberg, co-PI) and donors to The Alex Foundation.

References


Tommasi, L., & Vallortigara, G. (2001). Encoding of geometric and landmark information in the left and right hemispheres of the
Pepperberg


