



Quantifying the face inversion effect in nonhuman primates: a phylogenetic meta-analysis

Jason W. Griffin¹

Received: 29 July 2019 / Revised: 10 December 2019 / Accepted: 13 December 2019 / Published online: 23 December 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Face recognition is important for primate social cognition, enabling rapid discrimination between faces and objects. In humans, face recognition is characterized by certain cognitive specializations such as face-specific sensitivity to upright faces. The face inversion effect reflects the disproportionate inversion cost (i.e., poorer recognition of inverted compared to upright stimuli) for faces compared to non-face objects. Numerous studies have evaluated the face inversion effect in various nonhuman primates, but the findings are highly variable and mixed, especially in monkeys. To address this inconsistency, the current study employed a multilevel, phylogenetic meta-analysis on 52 effect sizes across 16 studies to quantify the magnitude of the face inversion effect in nonhuman primates. Overall, the difference in inversion costs for faces compared to nonface stimuli was small but not statistically significant ($b = 0.31$, $se = 0.16$, 95% CI $[-0.01, 0.62]$, $p = 0.06$). Additionally, this did not systematically vary between the family of species (chimpanzee or monkey) or type of face stimuli (conspecific or heterospecific). Ultimately, it cannot be concluded that the face inversion effect is a reliable phenomenon in either chimpanzees or monkeys, suggesting that nonhuman primates may not use similar face-specific processing strategies as humans.

Keywords Comparative · Face processing · Chimpanzee · Monkey · Face perception

Introduction

Face recognition is a highly adaptive cognitive ability integral for primate social cognition, because rapid recognition of friendly and non-friendly conspecific faces facilitates not only survival, but also successful social interaction among multiple group members (Griffin and Motta-Mena 2019). However, discriminating between faces is a computationally challenging ability, because faces are a highly homogenous class of visual stimuli that share the same basic features and configuration. Despite this, humans (*Homo sapiens*) can recognize thousands of individual faces accurately and quickly (Hsiao and Cottrell 2008). The ease with which humans can discriminate faces fundamentally depends on

holistic processing. Accordingly, faces are not perceived and represented by their individual facial features, but they are represented as a unified perceptual whole (for review, Tanaka and Simonyi 2015). Therefore, sensitivity to this basic configuration enables accurate and rapid recognition of subtle variations in the basic configuration among hundreds of unique faces.

Humans are remarkable at recognition for upright faces; however, when faces are presented upside down, this is marked by a large inversion cost, which is the decrease in accuracy and speed of recognition for inverted relative to upright stimuli (see meta-analysis, Bruyer 2011). Importantly, the inversion cost is much larger for faces compared to a wide range of other classes of visual stimuli including cars, scenes, and houses (Farah et al. 1995; Rossion 2008, 2009; Yin 1969). The disproportionate inversion cost, a phenomenon known as the face inversion effect (FIE), is due to inverted faces disrupting the canonical configuration to which holistic face-processing mechanisms are tuned. As a result, inverted faces cannot be processing holistically and elicit the use of feature-based recognition strategies that are similar to strategies used for non-face objects, but are less effective for face recognition (Farah et al. 1995; Rossion

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10071-019-01340-8>) contains supplementary material, which is available to authorized users.

✉ Jason W. Griffin
jxg569@psu.edu

¹ Department of Psychology, Pennsylvania State University, 423 Moore Building, University Park, PA 16802, USA

2008, 2009). Importantly, feature-based strategies for non-face objects are less sensitive to inversion and are more efficient for identifying non-face objects than faces. Taken together, this would indicate that faces are disproportionately dependent on holistic processing compared to non-face objects (Farah et al. 1995; Rossion 2008, 2009).

The development of the FIE is widely debated but has been dominated by two hypotheses: the *face-specific mechanism hypothesis* (Kanwisher 2000; McKone et al. 2007; Yin 1969) and the *expertise hypothesis* (Diamond and Carey 1986; Gauthier et al. 2000; Richler and Gauthier 2014). The latter suggests that the FIE is the result of face-specific cognitive and neural specializations over human and nonhuman primate evolutionary history. This stems from a number of human and nonhuman primate findings including the original FIE (Yin 1969), infant visual preferences to upright face-like information (Goren et al. 1975), and face-selective neural sensitivity to upright faces in the fusiform face area (Kanwisher et al. 1998; Yovel and Kanwisher 2005). In contrast, the *expertise hypothesis* suggests that face perception is enabled by mechanisms that operate on classes of stimuli for which people have developed expertise. For faces, most humans have developed exceptional expertise due to a lifetime of experience of perceiving and recognizing faces. However, the visual system can develop expertise for other types of stimuli that share a first-order configuration such as cars, chess boards, and birds (Boggan et al. 2012; Bukach et al. 2010; Gauthier et al. 2000). Importantly, both hypotheses contend that upright faces are processed holistically; the main difference is that the *expertise hypothesis* suggests this expert-like processing strategy is the result of massive exposure to faces, whereas the *face-selective mechanism hypothesis* suggests this processing strategy is due to innate, functionally specialized neural systems.

Face inversion effect in nonhuman primates

The FIE is one of the most widely studied face-processing phenomena in nonhuman primates including tamarins (*Saguinus oedipus*; Weiss et al. 2001), baboons (*Papio papio*; Parron and Fagot 2008), long-tailed macaques (*Macaca fascicularis*; Dittrich 1990), Japanese monkeys (*Macaca fuscata*; Tomonaga 1994), squirrel monkeys (*Saimiri sciureus*; Phelps and Roberts 1994), and an infant gibbon (*Hylobates agilis*; Myowa-Yamakoshi and Tomonaga 2001). However, the majority of research has focused on chimpanzees (*Pan troglodytes*) and rhesus monkeys (*Macaca macaque*) yielding highly variable and inconsistent results (for review, Parr 2011a; Rossion and Taubert 2019). For example, numerous articles have shown that chimpanzees show a reliable FIE (Dahl et al. 2013; Parr 2011b; Parr and Heintz 2005; Parr et al. 2006; Weldon et al. 2013; Wilson and Tomonaga 2018; Tomonaga 1999, 2007), although this finding is

not ubiquitous (Tomonaga et al. 1993; Kret and Tomonaga 2016; Gao and Tomonaga 2018). In rhesus monkeys, there is a more symmetrical distribution of studies supporting (Dahl et al. 2007; Neiwirth et al. 2006; Overman and Doty 1982; Tomonaga 1994; Vermeire and Hamilton 1998) and not supporting (Bruce 1982; Dittrich 1990; Gothard et al. 2004; Parr et al. 1999; Rosenfield and Van Hoesen 1979; Weiss, et al., 2001) the existence of the FIE. Based on the current state of the literature, there is little cumulative evidence to suggest that nonhuman primates show similar face inversion effects as humans and the magnitude of this effect has never been estimated across the existing literature. Aside from known limitations due to small sample sizes and imprecise estimates, this large inconsistency may be driven by the large methodological heterogeneity between studies.

Methodological heterogeneity

One of the core methodological differences that contributes to inconsistency in the literature is the numerous operationalizations of the FIE. In its original instantiation, the crucial aspect was that the inversion cost was much larger for faces than for other non-face stimuli (Yin 1969). Despite this, numerous studies in nonhuman primates have designed studies and made conclusions about the inversion cost between upright and inverted faces, without comparisons to nonface stimuli (Tomonaga 1994; Dahl et al. 2009; Guo et al. 2003). In contrast, a number of studies have reported the relative inversion costs between face and nonface stimuli, which reflects the appropriate interpretation of the effect (Rossion 2008, 2009). Regardless of definition, conclusions about the FIE from these studies are often compared directly, which contributes to the inconsistency in the existing literature (Parr 2011a).

In addition to differences in operationalization, the FIE is measured with various types of indirect and direct outcome measures in nonhuman primates. For example, a common indirect way of measuring recognition is through a visual preference paradigm called visual paired comparison (VPC), which leverages the fact that the visual system preferentially attends to novel stimuli. VPC is routinely used in nonverbal populations including infant humans and nonhuman primates since explicit recognition requires compliance with specific task instructions (e.g., “Select the face you remember”). In a face version of this task, subjects passively view a target face, then are subsequently shown the same face accompanied presented next to a novel face. Increased looking time to the novel face compared to the previously seen face (i.e., novelty preference) would be a positive indicator of face recognition. In contrary, face detection/categorization paradigms such as “Matching-to-sample” and “Go/No Go” are direct ways of measuring face recognition that require explicit memory of previously seen faces. Subjects

are shown a target face (i.e., a sample) during an encoding/study phase, which is followed by a two-alternative forced-choice paradigm where subjects must select the target face from a distractor face (i.e., match to the sample). This is the most commonly used method for evaluating face recognition in humans and nonhuman primates, because it directly evaluates an explicit form of recognition requiring the subject to actively make judgements about previously seen stimuli.

In addition to the general task paradigm, studies also differ in the types of stimuli used. For example, some studies exclusively focus on conspecific faces (e.g., Bruce 1982; Calcutt et al. 2017; Gao and Tomonaga 2018), whereas others focus on the inclusion of both conspecific and hetero-specific faces (e.g., Parron and Fagot 2008; Pokorny et al. 2011). Importantly, these studies have shown stimulus-specific effects. For example, Pokorny et al. (2011) showed inversion effects of conspecific and human faces, but not chimpanzee faces or cars in a group of capuchin monkeys. In chimpanzees, Parr et al. (1998) showed similar results suggesting inversion effects for conspecific and human faces, but not capuchin monkey faces or cars. More recently, Kret and Tomonaga (2016) also found similar inversion effects for conspecific and human faces in chimpanzees. This indicates that the type of face stimuli can have an impact on the presence of a FIE and may also moderate the magnitude of the FIE.

Synthesizing the existing literature

Primary research in comparative psychology is often limited by small sample sizes because of limited access to non-human primates. Although not exclusive to comparative psychology, small studies are vulnerable to high sampling error, overestimating effect sizes, generating estimates in the wrong direction, and replication failures (Camerer et al. 2018; Gelman and Carlin 2014). Fortunately, this issue can be addressed with the use of meta-analysis, a principled statistical procedure for aggregating data from multiple studies. Meta-analyses improves generalizability of results, confidence of published findings, and allows for a more precise effect size estimate than any individual study. In fact, without meta-analysis, it is nearly impossible to discern the direction, magnitude, and generalizability of observed effects from numerous published findings (Meehl 1954). Meta-analytic techniques can also provide information on publication bias, between-study variability (heterogeneity among studies), and moderation of effect size estimates due to sample- and study-related factors.

The fidelity of any meta-analysis is contingent on three fundamental principles including a precise operationalization of the effect under evaluation, a clear set of inclusion and exclusion criteria, and proper assessment of methodological and statistical heterogeneity. Failure to adhere to

these core principles can lead to the “garbage in, garbage out” phenomenon. Therefore, I first ensured consistency in outcome measures by operationalizing the FIE as the differential inversion cost for faces compared to non-face objects with the inversion cost being defined as the recognition disadvantage for inverted compared to upright stimuli. Although the FIE has been studied using face detection and visual preference paradigms, I only included “Matching-to-sample” and “Go/No Go” behavioral paradigms of recognition, because these are psychometrically related and the most commonly used paradigms in human and nonhuman primates. This also ensures consistency in outcome measures extracted from the existing literature. For example, I did not include visual preference paradigms, because it is widely argued that the novelty preference evaluates a different psychological construct than face detection tasks (implicit vs. explicit recognition) and these often do not show convergent validity (e.g., Pascalis et al. 2004). Finally, I accounted for and assessed relevant sources of heterogeneity with the use of a random-effects model and the evaluation of potential moderating variables.

The primary goal of this study is to estimate a summary effect size from the existing literature to answer three primary questions: (1) overall, is there a FIE in nonhuman primates? (2) If so, what is the magnitude of this effect? (3) Does the magnitude of this effect differ between species? Considering the large inconsistency of published findings, no specific predictions were made about the direction, magnitude, or moderation of face inversion effects in nonhuman primates. Although this study is not designed to explicitly test the face-selective mechanism or expertise account of the FIE, the current study can show support in favor of one or the other. For example, if the results show that the FIE is a statistically significant phenomenon in nonhuman primates, this would support the face-selective mechanism hypothesis indicating that face-selective sensitivity to upright faces is a common feature among all primates. However, this finding would not necessarily rule out the expertise account of the FIE, because it is unlikely that any of the animals have extensive exposure to the non-face objects used in the literature. However, failure to find any evidence of a FIE in nonhuman primates would lend support to the expertise hypothesis indicating that these animals need more exposure and experience with numerous faces to develop a similar holistic processing strategy as humans.

Methods

Study search and identification

A PubMed database search was conducted on July 3, 2019 (updated on November 8, 2019) with the following search

parameters: (“processing” OR “recognition” OR “recall” OR “perception” OR “memory”) AND (“primate” OR “monkey” OR “monkeys” OR “chimpanzee” OR “chimpanzees” OR “bonobo” OR “bonobos” OR “Rhesus” OR “troglodyte” OR “troglodytes” OR “Macaca” OR “macaque” OR “macaques”) AND (“face” OR “facial” OR “object” OR “shape” OR “scene” OR “car” OR “automobile” OR “house”) AND (“inversion” OR “upside down” OR “rotation” OR “inverted” OR “holistic” OR “configural”). In addition, the reference list from Parr (2011a) was appended to these search results. Finally, a manual reference search was conducted by screening the references lists of each article meeting full inclusion criteria from the initial search results. All Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines were followed including adherence to the PRISMA checklist (Supplemental Table 1) and flow diagram (Fig. 1; Moher et al. 2009).

Inclusion and exclusion criteria

Inclusion criteria

Articles must (1) include a sample of nonhuman primates; (2) report accuracy performance differences between upright and inverted stimuli (faces and/or objects) in a perception/recognition task; and (3) report standardized mean change scores between these conditions (upright, inverted) or enough information to calculate one (e.g., Cohen’s d , t , M , SD , N , se , raw individual data).

Exclusion criteria

Articles were excluded if they were (1) not published in English; (2) not peer reviewed; (3) not an empirical study. If articles met inclusion criteria in all categories but failed to include information regarding the effect size, I emailed corresponding authors to request the relevant data. If the data were not provided, the study was excluded.

Study selection

Database search

Inclusion criteria were applied in two steps. First, all Titles and Abstracts from the initial PubMed search were screened for inclusion. Since this was the first level of screening, over-inclusion was emphasized to maximize study yield. For example, abstracts were only excluded by demonstrating clear examples for exclusion such as not being published in English or not an empirical article (e.g., review or commentary). After this, Full Text articles were screened for

inclusion by evaluating the entire article for the primary inclusion components (see inclusion criteria).

Manual reference search

After all relevant articles were full-text reviewed, a manual search of each reference list was conducted on all articles meeting full inclusion criteria. To do this systematically, *Regex*, a string pattern recognition syntax, was used to extract article titles from each of the reference lists. Once these articles were extracted and duplicates were removed, the article titles were screened for inclusion. Then, I repeated the two steps described above: (1) Title and abstract screening; and (2) Full-text review.

Data extraction and computing effect sizes

The FIE was defined as the disproportionate inversion cost for faces compared to nonface objects. Therefore, within-subject accuracy performance for upright and inverted stimuli (faces and objects) was extracted to compute the standardized mean change score (Cohen’s d_z) for each object class (Cumming 2012). For articles only reporting M , SD , and N , the following set of formulas were used to compute the standardized mean change scores between upright and inverted conditions (see Cumming 2012; Viechtbauer 2010):

$$d_z = \frac{\bar{X}_2 - \bar{X}_1}{s_{diff}}, \quad (1)$$

where \bar{X}_2 is the upright mean score and \bar{X}_1 is the inverted mean score, and s_{diff} is the difference in standard deviation between scores calculated with:

$$s_{diff} = \sqrt{s_1^2 + s_2^2 - 2rs_1s_2}, \quad (2)$$

where r is the correlation between \bar{X}_2 and \bar{X}_1 . Since this correlation is rarely reported, a moderate correlation ($r=0.5$) was assumed for all studies and a sensitivity analysis was conducted with various correlation imputations between 0 and 1. The upright and inverted standard deviations are s_2 and s_1 . For studies that reported paired t values, these were converted directly to d_z (Cumming 2012; Lakens 2013):

$$d_z = \frac{t}{\sqrt{n}} \quad (3)$$

Once all effect sizes, d_z , were computed, the sample variances were computed with the following:

$$V[d_z] = \frac{1}{n} + \frac{d_z^2}{2n}. \quad (4)$$

Cohen’s d_z is known to overestimate the population effect size, especially when degrees of freedom are low ($df < 50$); therefore, d_z was converted to Hedge’s g , which corrects for

Table 1 Summary of the study and effect size characteristics, listed chronologically, for studies evaluating performance accuracy of upright and inverted stimuli in nonhuman primates

Study	Study#	ES #	Species	Stimuli	Stimuli type	<i>N</i>	<i>g</i>	<i>se</i>	Lower	Upper
Bruce (1982)[1]	1	1	<i>Macaca fascicularis</i>	Face	CS	6	0.05	0.41	-0.75	0.85
Bruce (1982)[2]	1	2	<i>Macaca fascicularis</i>	Face	CS	6	0.06	0.41	-0.74	0.86
Phelps and Roberts (1994)[1]	2	3	<i>Saimiri sciureus</i>	Face	HS	1	4.67	3.45	-2.09	11.43
Phelps and Roberts (1994)[2]	2	4	<i>Saimiri sciureus</i>	Nonface	scene	1	1.64	1.53	-1.36	4.64
Parr et al. (1998)[1]	3	5	<i>Pan troglodytes</i>	Face	CS	5	0.79	0.51	-0.22	1.79
Parr et al. (1998)[2]	3	6	<i>Pan troglodytes</i>	Face	HS	5	1.62	0.68	0.29	2.95
Parr et al. (1998)[3]	3	7	<i>Pan troglodytes</i>	Nonface	car	5	0.59	0.48	-0.36	1.54
Tomonaga (1999)[1]	4	8	<i>Pan troglodytes</i>	Face	HS	1	2.18	1.84	-1.42	5.78
Parr (1999)[1]	5	9	<i>Macaca mulatta</i>	Face	CS	4	2.67	1.07	0.58	4.77
Parr (1999)[2]	5	10	<i>Macaca mulatta</i>	Face	HS	4	1.19	0.65	-0.09	2.47
Parr (1999)[3]	5	11	<i>Macaca mulatta</i>	Face	HS	4	0.83	0.58	-0.31	1.96
Parr (1999)[4]	5	12	<i>Macaca mulatta</i>	Nonface	car	4	2.3	0.95	0.43	4.17
Parr (1999)[5]	5	13	<i>Macaca mulatta</i>	Nonface	shape	4	0.12	0.5	-0.87	1.1
Parr and Heintz (2005)[1]	6	14	<i>Pan troglodytes</i>	Face	CS	6	0.74	0.46	-0.17	1.64
Parr and Heintz (2005)[2]	6	15	<i>Pan troglodytes</i>	Nonface	house	6	-0.32	0.42	-1.14	0.5
Parron and Fagot (2008)[1]	7	16	<i>Papio papio</i>	Face	both	7	-0.03	0.38	-0.77	0.71
Parr and Heintz (2008)[1]	8	17	<i>Macaca mulatta</i>	Face	CS	5	0.98	0.54	-0.09	2.05
Parr and Heintz (2008)[2]	8	18	<i>Macaca mulatta</i>	Face	HS	5	1.06	0.56	-0.04	2.15
Parr and Heintz (2008)[3]	8	19	<i>Macaca mulatta</i>	Nonface	house	5	0.99	0.55	-0.08	2.06
Parr et al. (2008a)[1]	9	20	<i>Pan troglodytes</i>	Face	CS	6	1.17	0.53	0.13	2.21
Parr et al. (2008a)[2]	9	21	<i>Pan troglodytes</i>	Face	CS	6	1.34	0.56	0.24	2.45
Parr et al. (2008a)[3]	9	22	<i>Pan troglodytes</i>	Face	CS	6	2.04	0.72	0.63	3.44
Parr et al. (2008a)[4]	9	23	<i>Pan troglodytes</i>	Face	CS	6	3.13	0.99	1.18	5.07
Parr et al. (2008a)[5]	9	24	<i>Pan troglodytes</i>	Face	CS	6	1.17	0.53	0.13	2.21
Parr et al. (2008a)[6]	9	25	<i>Pan troglodytes</i>	Face	CS	6	1.4	0.57	0.28	2.53
Parr et al. (2008b)[1]	10	26	<i>Macaca mulatta</i>	Face	CS	8	1.14	0.45	0.25	2.04
Parr et al. (2008b)[2]	10	27	<i>Macaca mulatta</i>	Face	HS	8	0.81	0.41	0.01	1.61
Parr et al. (2008b)[3]	10	28	<i>Macaca mulatta</i>	Face	HS	8	0.97	0.43	0.13	1.81
Parr et al. (2008b)[4]	10	29	<i>Macaca mulatta</i>	Nonface	house	8	0.52	0.38	-0.22	1.26
Parr et al. (2008b)[5]	10	30	<i>Macaca mulatta</i>	Nonface	Clip art	8	0.69	0.39	-0.08	1.46
Parr (2011a, b)[1]	11	31	<i>Macaca mulatta</i>	Face	CS	6	0.91	0.49	-0.04	1.86
Parr (2011a, b)[2]	11	32	<i>Macaca mulatta</i>	Face	HS	6	1.43	0.58	0.29	2.57
Parr (2011a, b)[3]	11	33	<i>Pan troglodytes</i>	Face	HS	6	0.11	0.41	-0.7	0.91
Parr (2011a, b)[4]	11	34	<i>Pan troglodytes</i>	Face	CS	6	1.55	0.61	0.36	2.74
Parr (2011a, b)[5]	11	35	<i>Macaca mulatta</i>	Nonface	house	5	1.17	0.58	0.03	2.3
Parr (2011a, b)[6]	11	36	<i>Pan troglodytes</i>	Nonface	house	5	0.63	0.49	-0.33	1.59
Weldon et al. (2013)[1]	12	37	<i>Macaca mulatta</i>	Face	CS	6	0.05	0.41	-0.75	0.86
Weldon et al. (2013)[2]	12	38	<i>Pan troglodytes</i>	Face	CS	5	1.9	0.75	0.43	3.37
Kret and Tomonaga (2016)[1]	13	39	<i>Pan troglodytes</i>	Face	HS	5	0.06	0.45	-0.82	0.94
Kret and Tomonaga (2016)[2]	13	40	<i>Pan troglodytes</i>	Face	CS	5	-0.02	0.45	-0.89	0.86
Kret and Tomonaga (2016)[3]	13	41	<i>Pan troglodytes</i>	Face	HS*	5	0.04	0.45	-0.84	0.92
Kret and Tomonaga (2016)[4]	13	42	<i>Pan troglodytes</i>	Face	CS*	5	0.12	0.45	-0.76	1
Kret and Tomonaga (2016)[5]	13	43	<i>Pan troglodytes</i>	Nonface	CS foot	5	0.02	0.45	-0.86	0.89
Kret and Tomonaga (2016)[6]	13	44	<i>Pan troglodytes</i>	Nonface	HS foot	5	-0.04	0.45	-0.91	0.84
Kret and Tomonaga (2016)[7]	13	45	<i>Pan troglodytes</i>	Nonface	CS foot*	5	0.15	0.45	-0.74	1.03
Kret and Tomonaga (2016)[8]	13	46	<i>Pan troglodytes</i>	Nonface	Car*	5	-0.07	0.45	-0.94	0.81
Kret and Tomonaga (2016)[9]	13	47	<i>Pan troglodytes</i>	Nonface	HS foot*	5	0.03	0.45	-0.84	0.91
Calcutt et al. (2017)[1]	14	48	<i>Cebus apella</i>	Face	CS	9	2.77	0.73	1.33	4.21
Gao and Tomonaga (2018)[1]	15	49	<i>Pan troglodytes</i>	Face	CS	6	0.15	0.41	-0.65	0.96
Gao and Tomonaga (2018)[2]	15	50	<i>Pan troglodytes</i>	Nonface	House	6	-0.06	0.41	-0.86	0.74
Gao and Tomonaga (2018)[3]	15	51	<i>Pan troglodytes</i>	Nonface	House	6	0.17	0.41	-0.64	0.97
Wilson and Tomonaga (2018)[1]	16	52	<i>Pan troglodytes</i>	face	Both	5	1.09	0.57	-0.01	2.2

g standardized mean change score (Hedge's *g*), *se* standard error, *CS* conspecific stimuli, *HS* heterospecific stimuli

*Greyscale stimuli

this overestimation bias and can be interpreted exactly like Cohen's d_z (Cumming 2012; Hedges, 1981):

$$\text{Hedge's } g = \left(1 - \frac{3}{4df-1}\right) \times d_z. \quad (5)$$

Hierarchical data structure and sources of nonindependence (e.g., multiple effect sizes per study) are common in meta-analysis and methods of accounting for this multilevel structure have only been recently developed (Assink and Wibbelink 2016; Cheung 2014, 2019). Nonhuman primate research has a fairly unique source of nonindependence, which is reflected by the limited number of animals available for research. As such, the same individual animals are often used for multiple published studies and these studies often report multiple effect sizes. As such, each animal location (e.g., Yerkes National Primate Research Center), unique study, and each effect size was uniquely dummy coded to model this nested structure. For the moderation analyses, the sample species, stimuli category (face, object), and type of face stimuli (conspecific, heterospecific) corresponding to each effect size were extracted.

Phylogenetic relatedness

A phylogenetic tree was obtained from the primate 10KTree database by selecting the species represented in the final dataset and evaluating 100 trees for convergence (Arnold et al. 2010). Then, the relevant phylogenetic correlation matrix among species was extracted with 'ape' to specify the covariance structure of all subsequent models (Paradis and Schliep 2018).

Statistical analysis

Data were analyzed using multilevel meta-analysis via the 'metafor' package for the statistical software R (R Core Team 2018; Viechtbauer 2010). Given the methodological heterogeneity in reviewed studies, a random-effects, phylogenetic meta-analytic model was used to estimate the overall effect size difference between face and nonface object inversion costs, while accounting for dependencies in observations (e.g., multiple effect sizes per study; phylogenetic relatedness among species). Specifically, I fit a multilevel model to partition different sources of variance: sampling variance of the observed effect sizes, variance between effect sizes from the same study, and variance between studies. In addition to these sources of nonindependence, the covariance structure was specified explicitly based on the phylogenetic correlation matrix for species included in the analysis. This model was used to estimate the differences in inversion cost for faces and nonface objects, estimate within- and between-study variability, and account for relatedness among species.

This multilevel modeling approach, unlike traditional ones, properly handles dependency in observations, can include more effect sizes, achieves higher statistical power, and incorporates the phylogenetic relatedness among species (Cheung 2014; Lajeunesse 2009).

Sensitivity analysis

As shown in Eq. 2, the correlation between upright (\bar{X}_2) and inverted (X_1) conditions is required, but studies rarely report this correlation making it impossible to directly compute Cohen's d_z . As recommended by Viechtbauer (2010) and Borenstein et al. (2009), various estimates of r (i.e., 0.1, 0.3, 0.5, 0.7, 0.9) were inputted to evaluate if major changes in the summary estimates occurred with these systematic changes in the assumed correlation between the repeated measurements.

Moderation analysis

To evaluate effect size differences between monkeys and chimpanzees, the meta-analytic model was extended to include the moderator variable of species category where monkeys were defined by the following species: *Cebus apella*, *Macaca fascicularis*, *Macaca Mulatta*, *Saimiri sciureus*, and *Papio papio*. The chimpanzee category was comprised of *Pan troglodytes*. In addition, the moderator variable of face stimuli was defined as either conspecific if the faces matched the sample species, or heterospecific if the faces did not match the sample species.

Results

Study selection

Full description of study screening and identification is shown in Fig. 1. The initial study search (PubMed and Parr 2011a) yielded a total of 280 articles. After removing duplicates and conducting a title and abstract screening, 196 articles were excluded (see reasons in Fig. 1). The remaining 84 articles were full-text screened for inclusion of which 16 met full inclusion criteria. It should be noted that five additional articles met inclusion with the exception of providing the relevant information to compute an effect size. All corresponding authors were emailed for the relevant data; however, zero authors responded, and the studies had to be excluded. The reference lists of each of the 16 studies meeting full inclusion criteria ($n = 566$) were extracted, to conduct a manual reference search. After removing duplicates within the manual reference search and the initial search ($n = 277$), a title screening ruled out 247 nonrelevant articles (e.g., books, statistical packages). After another round

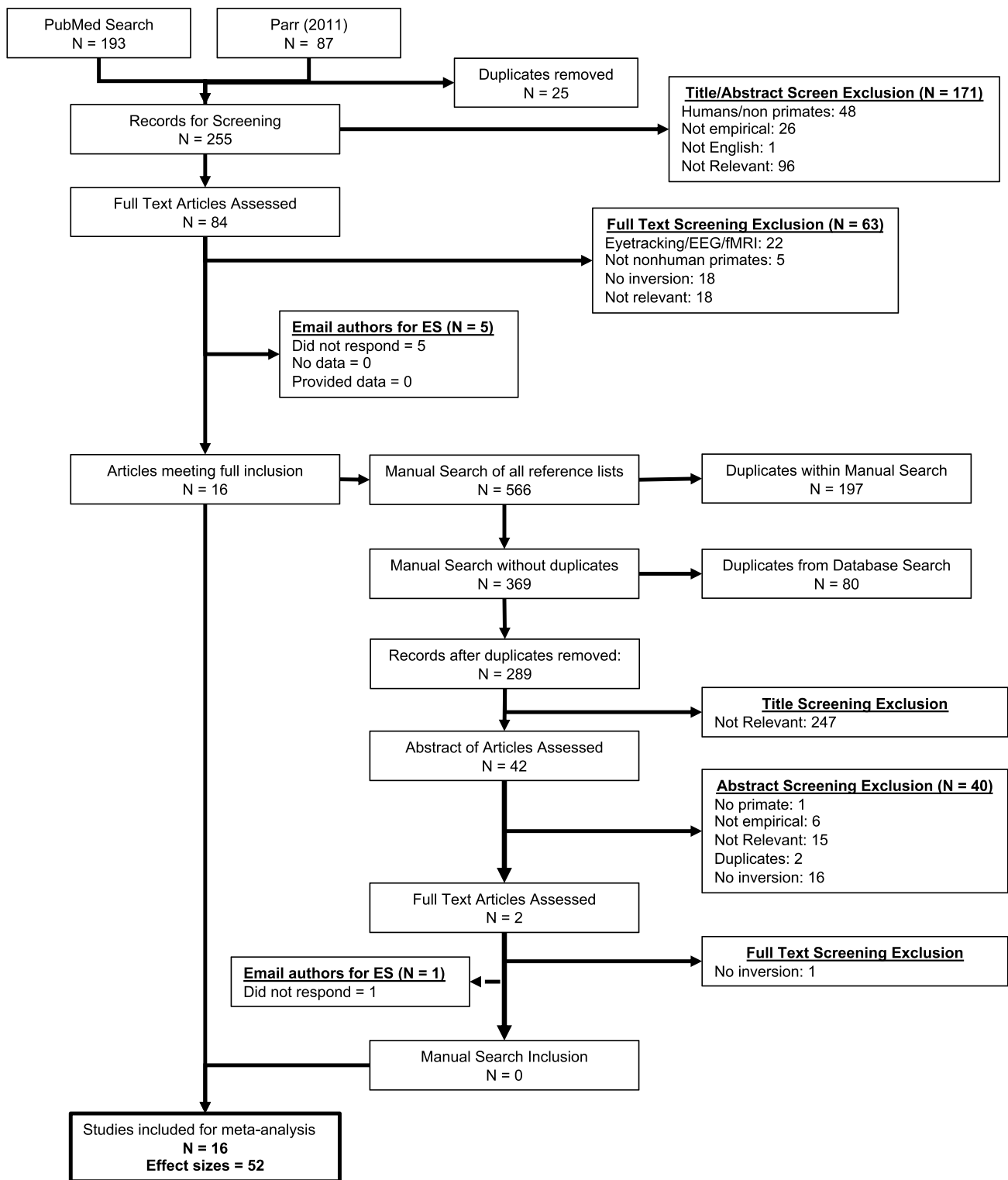


Fig. 1 PRISMA flow diagram of the article identification, screening, review process for the initial and manual search strategies with reasons for exclusion at each level

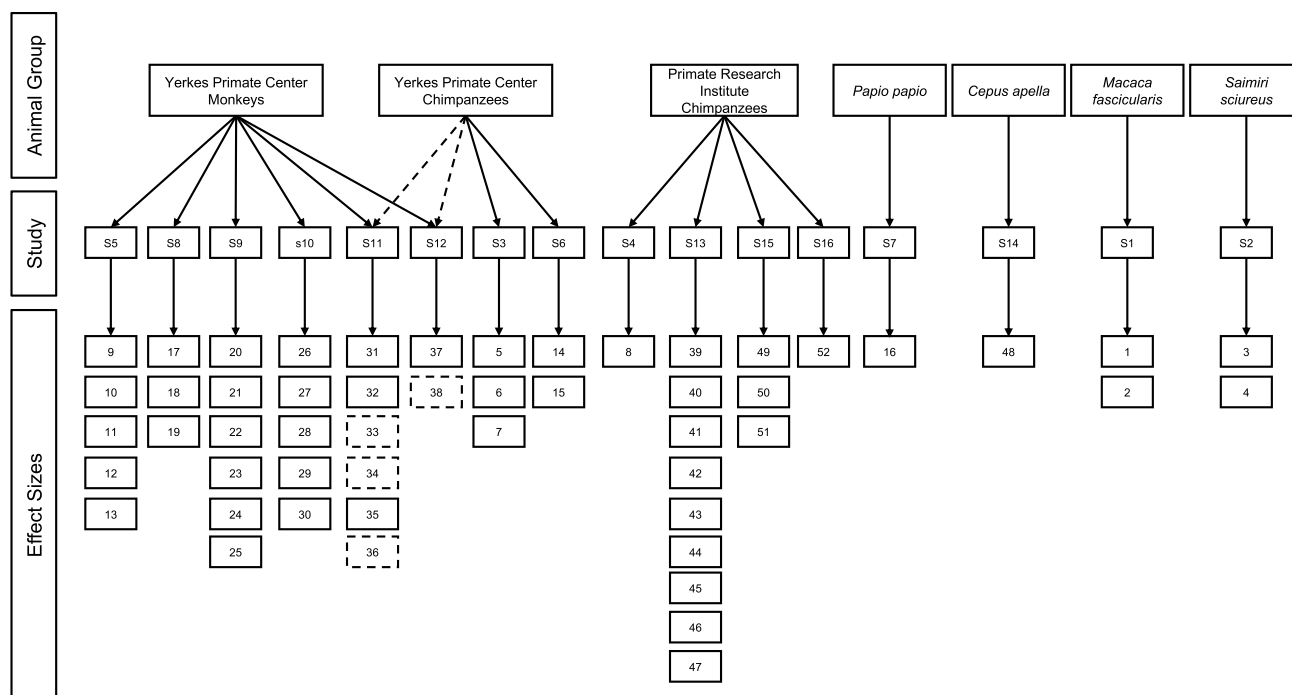


Fig. 2 Diagram of the multilevel structure inherent in the meta-analytic dataset. Individual effect sizes ($k=52$) were nested within study ($n=16$), which were nested within animal groups from various insti-

tutions. We specified this multilevel structure for all statistical models to account for nonindependence in the data. Dashed lines correspond to the specific effect sizes for chimpanzees (as opposed to Monkeys)

of title and abstracting screening, and full-text review, zero articles met inclusion criteria from this method.

In total, 52 effect sizes from 16 unique articles were included in the meta-analysis. This dataset included at least 43 unique nonhuman primates with representation from 6 different species (see Table 1). Since many of the animals came from the same laboratories and sites, the study with the largest sample size from that site (e.g., Yerkes National Primate Research Center) was assumed to be the total number of unique animals (see Fig. 2 for multilevel structure). The sample sizes, means, standard deviations, effect sizes, and species names are presented in Table 1.

After selecting the 6 species represented in the final dataset, all 100 phylogenetic tree estimates from 10Ktree converged to the same relatedness structure (see Supplemental Materials). The extracted correlation matrix was used to specify the covariance structure in all subsequent models.

Summary effect size differences

The overall effect size differences in inversion cost was larger for faces than for nonface objects ($b=0.31$, $se=0.16$, 95% CI $[-0.01, 0.62]$, $p=0.06$); however, this failed to reach statistical significance (see Fig. 3). Consistent with the choice of a random-effects model, there was large heterogeneity among the included effect sizes ($Q=78.56$,

$p=0.006$, $I^2=77.85\%$). Furthermore, the phylogenetic multilevel model revealed that this heterogeneity was due to the phylogenetic relationship among species (61.63%), differences between animal locations (8.00%), and between-study variation (8.29%). In contrast, there was virtually no within-study variability (0.00%). Since only 16 unique studies were obtained from the systematic search, “small study effects”, like publication bias, could not be evaluated, because all methods of assessment are unreliable when the number of studies is low (Debray et al. 2018).

Simple main effects showed that the summary effect size was large for faces (Hedge’s $g=1.13$, $se=0.59$, 95% CI $[-0.05, 2.31]$, $p=0.06$) and objects (Hedge’s $g=0.82$, $se=0.59$, 95% CI $[-0.38, 2.03]$, $p=0.17$), but both failed to research statistical significance. Regardless, the summary effect size for faces was larger than objects, but estimated imprecisely.

Effect of species category and face stimuli (moderation analysis)

One of the core goals of this study was to evaluate the extent to which monkeys and chimpanzees may differ in either the presence or differential magnitude of the FIE. Consistent with the overall model, the FIE was not moderated by species category such that monkey and chimpanzee inversion

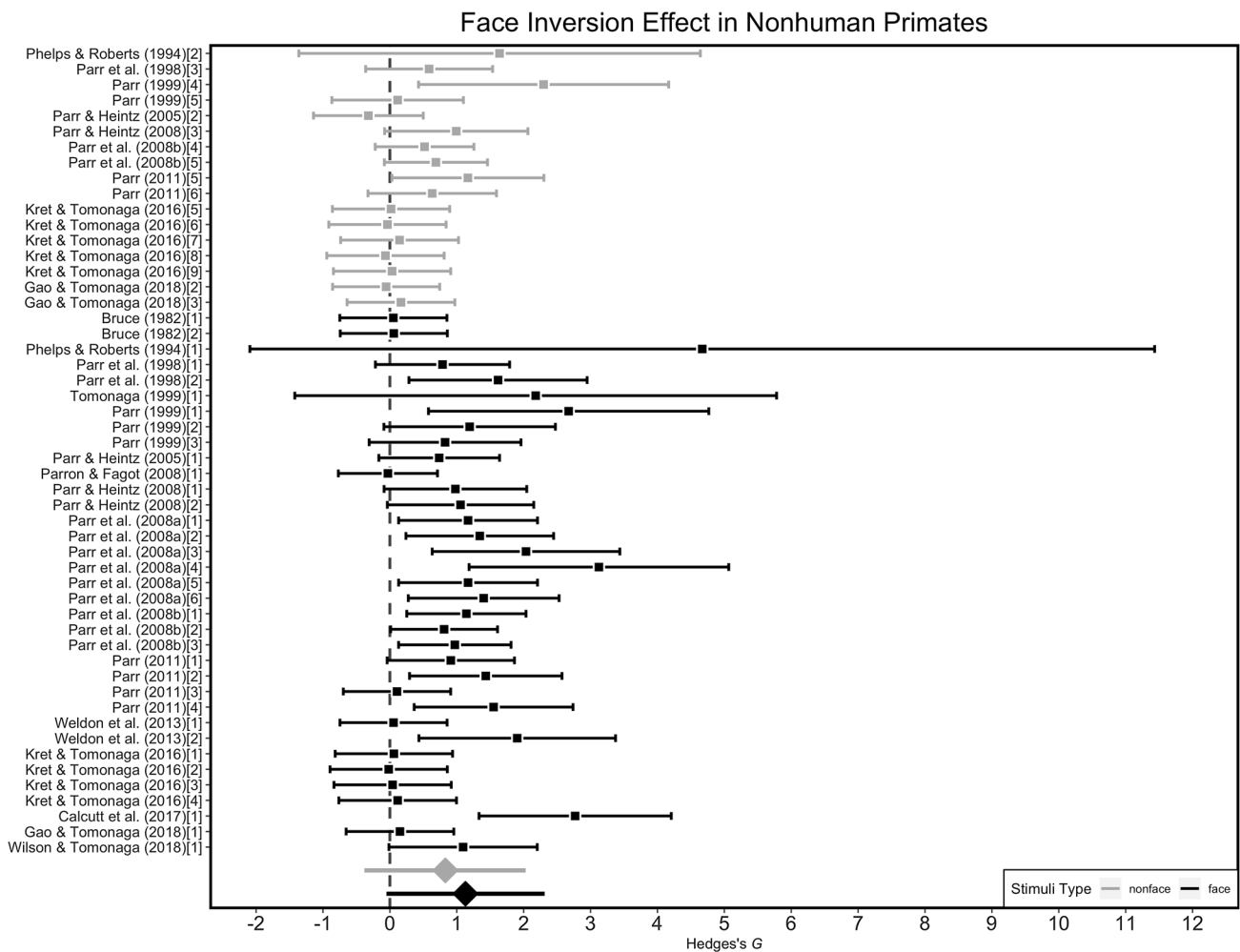


Fig. 3 Multilevel forest plot displaying the standardized mean change scores (Hedge’s *g*) for each study with effect sizes nested within study (indicated by bracket). The solid black interval at the bottom reflects the overall summary effect for faces (Hedge’s *g* = 1.13, *se* = 0.59, 95%

CI [−0.05, 2.31], *p* = 0.06). The solid grey interval reflects the overall summary effect for nonface objects (Hedge’s *g* = 0.82, *se* = 0.59, 95% CI [−0.38, 2.03], *p* = 0.17). Center diamonds represent the point estimates with the line interval representing 95% confidence intervals

costs were comparable (*b* = −0.17, *se* = 0.32, 95% CI [0.81, 0.48], *p* = 0.61). Similarly, the FIE was not moderated by the type of face stimuli such that the magnitude was not different between conspecific and heterospecific faces (*b* = 0.09, *se* = 0.21, 95% CI [−0.33, 0.52], *p* = 0.67). Two studies were excluded when evaluating this moderator variable since effect sizes were averaged across conspecific and heterospecific faces (see Table 1).

Sensitivity analysis

For 32 of the effect sizes, the standardized mean change score (Cohen’s *d_z*) was computed directly from paired *t* values. However, for 20 of the 52 effect sizes (38%), Cohen’s *d_z* was computed from the *M*, *SD*, and *N* while inputting various correlation values (see Eq. 2). The overall summary effect

size between faces and nonface objects was robust to differences in these various correlation imputations (*r*_{0.1} = 0.29, *r*_{0.3} = 0.30, *r*_{0.5} = 0.30, *r*_{0.7} = 0.32, *r*_{0.9} = 0.35) indicating that regardless of the correlation between upright and inverted conditions, the effect remained relatively unaltered.

Discussion

This is the first meta-analysis quantitatively synthesizing the FIE across the existing nonhuman primate literature. The primary goals were to determine if there was a reliable FIE in nonhuman primates, quantify the magnitude of this effect, and evaluate if this effect was different in chimpanzees and monkeys. This meta-analysis revealed that when data were aggregated for at least 43 unique subjects,

nonhuman primates showed a larger inversion cost for faces (Hedge's $g = 1.13$) compared to nonface objects (Hedge's $g = 0.82$); however, this failed to reach statistical significance ($p = 0.06$). Importantly, neither the presence nor magnitude of this effect, was moderated by species category (chimpanzee, monkey) or type of face stimuli (conspecific, heterospecific). While it appears that the faces generate a higher inversion cost than objects, this phenomenon was not statistically significant.

Considerable debate has centered on whether chimpanzees and/or monkeys show reliable face inversion effects similar to that of humans (Parr 2011a; Rossion and Taubert 2019). For example, previous reviews of this literature have concluded that the FIE is reliably found in chimpanzees and inconsistently observed in monkeys (Parr 2011a); however, this interpretation of the existing literature was not ubiquitous (Rossion and Taubert 2019), and is not supported by the current meta-analysis. Without a quantitative meta-analysis, it is unclear how one would combine, weight, and assign the appropriate beta weights for each related study that differ on a number of methodological factors in a way that is objective (Meehl 1954). As such, numerous speculations have been generated to explain why chimpanzees, but not monkeys, show face inversion effects. Much of this discussion has been centered on why these different primates would have qualitatively different face-processing strategies. Despite these hypotheses, the data presented in this meta-analysis suggest nonhuman primates do not reliably suffer a disproportional inversion cost for faces compared to other objects. Furthermore, there is no strong evidence that chimpanzees and monkeys have different inversion costs for faces relative to non-face objects suggesting that face-specific sensitivity may not be the result of socio-ecology, visual bias to conspecific faces, or phylogenetic relationships. Therefore, primary studies showing the presence of face inversion effects in chimpanzees or monkeys are likely study-specific findings that may be driven by sampling error or lack of an appropriate control condition.

Small sample sizes mark an obvious design limitation in most nonhuman primate research. As such, parameter estimates are extremely variable and may be in the wrong direction (Gelman and Carlin 2014; Ioannidis 2005). Of the studies reviewed here, sample sizes ranged from 1 to 9. While huge barriers preclude the collection of large nonhuman primate samples, the research community should be cautious when making strong conclusions about neural and cognitive mechanisms in chimpanzees and monkeys. As the data presented here show the host of speculations as to why chimpanzees, but not monkeys, show inversion effects may be informed by spurious findings. One possible solution to this small sample problem is the utilization of prospective meta-analysis. This approach, unlike the retrospective strategy used here, involves groups of researchers at different

sites or institutions conducting multi-center studies in which methodology heterogeneity can be controlled and sampling error reduced (i.e., larger and more generalizable samples). Another solution lies in the use of Bayesian statistics. Specifically, the incorporation of prior knowledge (previous studies) and the ability to update the current state of knowledge based on new data (future studies). This inference strategy has become more common in ecology and facilitates a cumulative science in which future studies can directly build from previous studies. This is particularly helpful in primate research, when the availability of large samples is limited or even impossible to obtain.

The current study was not designed to directly assess the *face-selective mechanism hypothesis* or *expertise hypothesis* related to the development of holistic processing. However, the absence of a FIE in nonhuman primates after quantifying the existing literature using meta-analysis has clear implications for the two dominant hypotheses related to development of holistic processing. For example, the *face-selective mechanism hypothesis* suggests that face-specific holistic processing is an innate feature of primate face processing due to the importance and social relevance of faces in evolutionary history. However, this meta-analysis does not support such account of holistic processing since there was no evidence of a FIE in either chimpanzees or monkeys. If such face-specific mechanisms were in place at birth, there should have been a clear FIE at least in chimpanzees. In contrary, advocates of the *expertise hypothesis* would argue that the absence of a FIE is due to the lack of exposure to a large amount of faces and non-face objects. Furthermore, this may be the result of nonhuman primates routinely reared in unnatural settings altering the development of face-specific cognitive abilities like holistic processing.

Even though this meta-analysis represents the largest aggregation of data relevant to the FIE in nonhuman primates, all the animals were raised in Barren, Institutional, Zoo, and other Rare Rearing Environments (BIZARRE; Leavens et al. 2010). As such, the extent to which the results reported in this meta-analysis generalize to the population of chimpanzees and monkeys is unknown. For example, there are a number of studies that show that institutionalized nonhuman primates that are raised in unnatural environments can develop very different visual perceptual and communicative abilities than naturally dwelling chimpanzees (for review, Leavens and Racine 2009; Leavens et al. 2009). Considering this and the fact that nonhuman primates are usually trained to perform a specific task (e.g., face detection task), there is no reason to think that the results presented in this meta-analysis would generalize to wild nonhuman primates as a whole. However, the failure to find a statistically significant FIE in nonhuman primates suggests that at least for BIZARRE nonhuman primates, the FIE is not reliable

phenomenon. Future research can make great contributions to the field by developing more ecologically valid methods of face perception that can be generalized to the population of nonhuman primates as a whole.

This study was limited by its input data. For example, while the FIE is a commonly studied phenomenon in nonhuman primates, this literature is inherently small making it difficult to derive precise estimates. In addition, this limitation was accentuated, because five studies met full inclusion criteria but did not report relevant data. Upon email request, zero authors responded to the data request. This would have increased the total number of studies by 31%, providing a more precise estimate of the FIE. In addition, due to the limited number of studies, assessment of “small study effects”, which can include publication bias, was not possible. Finally, all the studies published on this topic are methodologically unique. That is, not one single study did the exact same thing as another. As a result, the FIE was evaluated using a variety of paradigms including matching-to-sample, delayed matching-to-sample, visual paired comparison, Go/NoGo, and the oddity task. To reduce the between-study heterogeneity, only studies that used a matching-to-sample, oddity paradigm, or a Go/NoGo task were included since these all measured explicit recognition performance. However, the exclusion of visual preference paradigms limits the number of studies reviewed here.

Conclusion

This is the first quantitative synthesis of the commonly studied FIE in nonhuman primates. The existing literature has been widely inconsistent on whether nonhuman primates show a FIE similar to humans. This meta-analysis addresses the apparent inconsistencies in the literature by pooling data across numerous studies and estimating the magnitude of the FIE in nonhuman primates. The differential inversion cost for faces compared to non-face objects was small, but not statistically distinguishable from each other. Importantly, this effect was not different between chimpanzees and monkeys or was it influenced by conspecific or heterospecific face stimuli. The failure to find reliable evidence of a FIE suggests that nonhuman primates may not have similar face-specific sensitivity to upright faces as humans.

Acknowledgements The author would like to thank Giorgia Picci for her helpful comments on the manuscript.

References

- Arnold C, Matthews LJ, Nunn CL (2010) The *10kTrees* website: a new online resource for primate phylogeny. *Evol Anthropol* 19:114–118
- Assink M, Wibbelink CJ (2016) Fitting three-level meta-analytic models in R: a step-by-step tutorial. *Quant Methods Psychol* 12(3):154–174. <https://doi.org/10.20982/tqmp.12.3.p154>
- Boggan AL, Bartlett JC, Krawczyk DC (2012) Chess masters show a hallmark of face processing with chess. *J Exp Psychol Gen* 141(1):37. <https://doi.org/10.1037/a0024236>
- Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2009) *Effect sizes based on means introduction to meta-analysis*. John Wiley & Sons Ltd, Chichester, pp 21–32. Doi: 10.1002/9780470743386.ch6.
- Bruce C (1982) Face recognition by monkeys: absence of an inversion effect. *Neuropsychologia* 20(5):515–521. [https://doi.org/10.1016/0028-3932\(82\)90025-2](https://doi.org/10.1016/0028-3932(82)90025-2)
- Bruyer R (2011) Configural face processing: a meta-analytic survey. *Perception* 40(12):1478–1490. <https://doi.org/10.1068/p6928>
- Bukach CM, Phillips SW, Gauthier I (2010) Limits of generalization between categories and implications for theories of category specificity. *Atten Percep Psychophys* 72(7):1865–1874. <https://doi.org/10.3758/app.72.7.1865>
- Calcutt SE, Rubin TL, Pokorny JJ, de Waal FB (2017) Discrimination of emotional facial expressions by tufted capuchin monkeys (*Sapajus apella*). *J Comp Psychol* 131(1):40. <https://doi.org/10.1037/com0000055>
- Camerer CF, Dreber A, Holzmeister F, Ho T-H, Huber J, Johannesson M, Wu H (2018) Evaluating the replicability of social science experiments in nature and science between 2010 and 2015. *Nature Human Behav* 2(9):637–644. <https://doi.org/10.1038/s41562-018-0399-z>
- Cheung M (2014) Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychol Methods* 19(2):211. <https://doi.org/10.1037/a0032968>
- Cheung M (2019) A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychol Rev*. <https://doi.org/10.1007/s11065-019-09415-6>
- Cumming G (2012) *Understanding the new statistics: effect sizes, confidence intervals, and meta-analysis*. Routledge, Abingdon
- Dahl CD, Logothetis NK, Hoffman KL (2007) Individuation and holistic processing of faces in rhesus monkeys. *Proc R Soc B Biol Sci* 274(1622):2069–2076. <https://doi.org/10.1098/rspb.2007.0477>
- Dahl CD, Wallraven C, Bühlhoff HH, Logothetis NK (2009) Humans and macaques employ similar face-processing strategies. *Curr Biol* 19(6):509–513. <https://doi.org/10.1016/j.cub.2009.01.061>
- Dahl CD, Rasch MJ, Tomonaga M, Adachi I (2013) The face inversion effect in non-human primates revisited—an investigation in chimpanzees (*Pan troglodytes*). *Sci Rep* 3(1):2504. <https://doi.org/10.1038/srep02504>
- Debray TP, Moons KG, Riley RD (2018) Detecting small-study effects and funnel plot asymmetry in meta-analysis of survival data: a comparison of new and existing tests. *Res Synth Methods* 9(1):41–50. <https://doi.org/10.1002/jrsm.1266>
- Diamond R, Carey S (1986) Why faces are and are not special: an effect of expertise. *J Exp Psychol Gen* 115(2):107–117. <https://doi.org/10.1037/0096-3445.115.2.107>
- Dittrich W (1990) Representation of faces in long-tailed macaques (*Macaca fascicularis*). *Ethology* 85(4):265–278. <https://doi.org/10.1111/j.1439-0310.1990.tb00406.x>
- Farah MJ, Tanaka JW, Drain MH (1995) What causes the face inversion effect? *J Exp Psychol Hum Percept Perform* 21(3):628. <https://doi.org/10.1037/0096-1523.21.3.628>
- Gao J, Tomonaga M (2018) The body inversion effect in chimpanzees (*Pan troglodytes*). *PLoS ONE* 13(10):e0204131. <https://doi.org/10.1371/journal.pone.0204131>
- Gauthier I, Skudlarski P, Gore JC, Anderson AW (2000) Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci* 3(2):191. <https://doi.org/10.1038/72140>

- Gelman A, Carlin J (2014) Beyond power calculations. *Perspect Psychol Sci* 9(6):641–651. <https://doi.org/10.1177/1745691614551642>
- Goren C, Sarty, Wu P (1975) Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics* 56(4):544–549
- Gothard KM, Erickson CA, Amaral DG (2004) How do rhesus monkeys (*Macaca mulatta*) scan faces in a visual paired comparison task? *Anim Cogn* 7(1):25–36. <https://doi.org/10.1007/s10071-003-0179-6>
- Griffin JW, Motta-Mena NV (2019) Face and object recognition. In: Shackelford T, Weekes-Shackelford VA (Eds) *Encyclopedia of evolutionary psychological science*. Springer, Cham.
- Guo K, Robertson RG, Mahmoodi S, Tadmor Y, Young MP (2003) How do monkeys view faces?—a study of eye movements. *Exp Brain Res* 150(3):363–374. <https://doi.org/10.1007/s00221-003-1429-1>
- Hedges LV (1981) Distribution theory for glass's estimator of effect size and related estimators. *J Educ Stat* 6(2):107–128. <https://doi.org/10.3102/10769986006002107>
- Hsiao J, Cottrell G (2008) Two fixations suffice in face recognition. *Psychol Sci* 19(10):998–1006. <https://doi.org/10.1111/j.1467-9280.2008.02191.x>
- Ioannidis JP (2005) Why most published research findings are false. *PLoS Med* 2(8):e124. <https://doi.org/10.1371/journal.pmed.0020124>
- Kanwisher N (2000) Domain specificity in face perception. *Nat Neurosci* 3(8):759–763. <https://doi.org/10.1038/77664>
- Kanwisher N, Tong F, Nakayama K (1998) The effect of face inversion on the human fusiform face area. *Cognition* 68(1):B1–B11. [https://doi.org/10.1016/S0010-0277\(98\)00035-3](https://doi.org/10.1016/S0010-0277(98)00035-3)
- Kret ME, Tomonaga M (2016) Getting to the bottom of face processing: species-specific inversion effects for faces and behinds in humans and chimpanzees (*Pan troglodytes*). *PLoS ONE* 11(11):e0165357. <https://doi.org/10.1371/journal.pone.0165357>
- Lajeunesse MJ (2009) Meta-analysis and the comparative phylogenetic method. *Am Nat* 174(3):369–381. <https://doi.org/10.1086/603628>
- Lakens D (2013) Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol* 4:863. <https://doi.org/10.3389/fpsyg.2013.00863>
- Leavens DA, Racine TP (2009) Joint attention in apes and humans: are humans unique? *J Conscious Stud* 16:240–267
- Leavens DA, Racine TP, Hopkins WD (2009) The ontogeny and phylogeny of non-verbal deixis. In: Botha R, Knight C (eds) *The prehistory of language*. Oxford University Press, Oxford, pp 142–165
- Leavens DA, Bard KA, Hopkins WD (2010) BIZARRE chimpanzees do not represent “the chimpanzee.” *Behav Brain Sci* 33:100–101. [Commentary on Henrich et al.]
- McKone E, Kanwisher N, Duchaine BC (2007) Can generic expertise explain special processing for faces? *Trends Cognit Sci* 11(1):8–15. <https://doi.org/10.1016/j.tics.2006.11.002>
- Meehl PE (1954) *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press, Minneapolis
- Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 6(7):e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Myowa-Yamakoshi M, Tomonaga M (2001) Development of face recognition in an infant gibbon (*Hylobates agilis*). *Infant Behav Dev* 24(2):215–227. [https://doi.org/10.1016/S0163-6383\(01\)00076-5](https://doi.org/10.1016/S0163-6383(01)00076-5)
- Neiwirth JJ, Hassett JM, Sylvester CJ (2006) Face processing in humans and new world monkeys: the influence of experiential and ecological factors. *Anim Cogn* 10(2):125. <https://doi.org/10.1007/s10071-006-0045-4>
- Overman WH, Doty RW (1982) Hemispheric specialization displayed by man but not macaques for analysis of faces. *Neuropsychologia* 20(2):113–128. [https://doi.org/10.1016/0028-3932\(82\)90002-1](https://doi.org/10.1016/0028-3932(82)90002-1)
- Paradis E, Schliep K (2018) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatic* 35:526–528
- Parr LA (2011a) The evolution of face processing in primates. *Philos Trans R Soc Lond Ser B Biol Sci* 366(1571):1764–1777. <https://doi.org/10.1098/rstb.2010.0358>
- Parr LA (2011b) The inversion effect reveals species differences in face processing. *Acta Physiol (Oxf)* 138(1):204–210. <https://doi.org/10.1016/j.actpsy.2011.06.005>
- Parr LA, Heintz M (2005) The perception of unfamiliar faces and houses by chimpanzees: influence of rotation angle. *Perception* 35(11):1473–1483. <https://doi.org/10.1068/p5455>
- Parr LA, Heintz M (2008) Discrimination of faces and houses by rhesus monkeys: the role of stimulus expertise and rotation angle. *Anim Cogn* 11(3):467–474. <https://doi.org/10.1007/s10071-008-0137-4>
- Parr LA, Winslow JT, Hopkins WD (1999) Is the inversion effect in rhesus monkeys face-specific? *Anim Cogn* 2(3):123–129. <https://doi.org/10.1007/s100710050032>
- Parr LA, Dove T, Hopkins WD (2006) Why faces may be special: evidence of the inversion effect in chimpanzees. *J Cognit Neurosci* 10(5):615–622. <https://doi.org/10.1162/089892998563013>
- Parr LA, Heintz M, Pradhan G (2008a) Rhesus monkeys (*Macaca mulatta*) lack expertise in face processing. *J Comp Psychol* 122(4):390. <https://doi.org/10.1037/0735-7036.122.4.390>
- Parr LA, Waller BM, Heintz M (2008b) Facial expression categorization by chimpanzees using standardized stimuli. *Emotion* 8(2):216. <https://doi.org/10.1037/1528-3542.8.2.216>
- Parron C, Fagot J (2008) Baboons (*Papio papio*) spontaneously process the first-order but not second-order configural properties of faces. *Am J Primatol* 70(5):415–422. <https://doi.org/10.1002/ajp.20503>
- Pascalis O, Hunkin NM, Holdstock JS, Isaac CL, Mayes AR (2004) Visual paired comparison performance is impaired in a patient with selective hippocampal lesions and relatively intact item recognition. *Neuropsychologia* 42(10):1293–1300. <https://doi.org/10.1016/j.neuropsychologia.2004.03.005>
- Phelps MT, Roberts WA (1994) Memory for pictures of upright and inverted primate faces in humans (*Homo sapiens*), squirrel monkeys (*Saimiri sciureus*), and pigeons (*Columba livia*). *J Comp Psychol* 108(2):114. <https://doi.org/10.1037/0735-7036.108.2.114>
- Pokorny JJ, Webb CE, de Waal FB (2011) An inversion effect modified by expertise in capuchin monkeys. *Anim Cogn* 14(6):839. <https://doi.org/10.1007/s10071-011-0417-2>
- R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Richler JJ, Gauthier I (2014) A meta-analysis and review of holistic face processing. *Psychol Bull* 140(5):1281. <https://doi.org/10.1037/a0037004>
- Rosenfeld SA, Van Hoesen GW (1979) Face recognition in the rhesus monkey. *Neuropsychologia* 17(5):503–509. [https://doi.org/10.1016/0028-3932\(79\)90057-5](https://doi.org/10.1016/0028-3932(79)90057-5)
- Rossion B (2008) Picture-plane inversion leads to qualitative changes of face perception. *Acta Physiol (Oxf)* 128(2):274–289. <https://doi.org/10.1016/j.actpsy.2008.02.003>
- Rossion B (2009) Distinguishing the cause and consequence of face inversion: the perceptual field hypothesis. *Acta Physiol (Oxf)* 132(3):300–312. <https://doi.org/10.1016/j.actpsy.2009.08.002>
- Rossion B, Taubert J (2019) What can we learn about human individual face recognition from experimental studies in monkeys? *Vis Res* 157:142–158. <https://doi.org/10.1016/j.visres.2018.03.012>
- Tanaka JW, Simonyi D (2015) The “parts and wholes” of face recognition: a review of the literature. *Q J Exp Psychol* 69(10):1876–1889. <https://doi.org/10.1080/17470218.2016.1146780>

- Tomonaga M (1994) How laboratory-raised Japanese monkeys (*Macaca fuscata*) perceive rotated photographs of monkeys: Evidence for an inversion effect in face perception. *Primates* 35(2):155–165. <https://doi.org/10.1007/bf02382051>
- Tomonaga M (1999) Inversion effect in perception of human faces in a chimpanzee (*Pan troglodytes*). *Primates* 40(3):417–438. <https://doi.org/10.1007/bf02557579>
- Tomonaga M (2007) Visual search for orientation of faces by a chimpanzee (*Pan troglodytes*): face-specific upright superiority and the role of facial configural properties. *Primates* 48(1):1–12. <https://doi.org/10.1007/s10329-006-0011-4>
- Tomonaga M, Itakura S, Matsuzawa T (1993) Superiority of conspecific faces and reduced inversion effect in face perception by a chimpanzee. *Folia Primatol* 61(2):110–114. <https://doi.org/10.1159/000156737>
- Vermeire BA, Hamilton CR (1998) Inversion effect for faces in split-brain monkeys. *Neuropsychologia* 36(10):1003–1014. [https://doi.org/10.1016/s0028-3932\(98\)00054-2](https://doi.org/10.1016/s0028-3932(98)00054-2)
- Viechtbauer W (2010) Conducting meta-analyses in R with the metafor package. *J Stat Softw* 36(3):1–48
- Weiss DJ, Kralik JD, Hauser MD (2001) Face processing in cotton-top tamarins (*Saguinus oedipus*). *Anim Cogn* 3(4):191–205. <https://doi.org/10.1007/s100710000076>
- Weldon KB, Taubert J, Smith CL, Parr LA (2013) How the Thatcher illusion reveals evolutionary differences in the face processing of primates. *Anim Cogn* 16(5):691–700. <https://doi.org/10.1007/s10071-013-0604-4>
- Wilson DA, Tomonaga M (2018) Visual discrimination of primate species based on faces in chimpanzees. *Primates* 59(3):243–251. <https://doi.org/10.1007/s10329-018-0649-8>
- Yin RK (1969) Looking at upside-down faces. *J Exp Psychol* 81(1):141–145. <https://doi.org/10.1037/h0027474>
- Yovel G, Kanwisher N (2005) The neural basis of the behavioral face-inversion effect. *Curr Biol* 15(24):2256–2262. <https://doi.org/10.1016/j.cub.2005.10.072>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.