Evolutionary Theories of Human Facial Attractiveness: Computer Graphic Investigations

Ian Penton-Voak

March, 2000

A thesis submitted in Septmeber, 1999, to the University of St. Andrews for the degree of Doctor of Philosophy in the School of Psychology

Declarations:

I, Ian Scott Penton-Voak, hereby certify that this thesis, which is approximately 49,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree

Date... 23-3-00 .... .signature of candidate... .

I was admitted as a candidate for the degree of Doctor of Philosophy in October 1996; the higher study for which this is a record was carried out in the University of St Andrews between 1996 and 1999.

Date... 23.3.00 .... .signature of candidate... ...

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

Date... 23.3.00 ....signature of supervisor..

Thesis copyright:

Unrestricted

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

Date… … … … … … ..signature of candidate… … … … … … … … … … … … … … … ..

Acknowledgements

I would like to thank the many people who have helped me produce this thesis. I have benefited enormously from the fantastic supervision, enthusiasm and generosity of Dave Perrett over the last three years. Several people have helped with the running of some of the experiments reported here. Thanks go to Duncan Castles, for help with the Japanese based work; Lindsey Murray & Karen Brandt, for kind assistance in data collection for the UK based menstrual cycle research. Thanks also to Jon Pierce for collecting much of the data in experiment 6. The computing skills of Duncan Rowland & Mike Burt have been and continue to be invaluable. Anne Perrett's proofreading of manuscripts has been much appreciated. The perception lab is a fun place to work, so thanks to everyone in the group, past and present, for making it so. Outside of the perception lab, Andy Whiten, Dick Byrne, Jeff Graves, Don Symons and Randy Thornhill have given advice, comments and constructive criticism on the work reported here.

Away from work, thanks to Rich & Christine, David & Jane, Kevin, Lindsey, Ted, Chris, Nathan, Clare, Paul, Tom and Kenny for their friendship and favours too diverse to mention. Finally, thanks to my parents, and most importantly, Sophie.

Contents

List of Figures:

Abstract

Evolutionary psychological approaches to facial attractiveness suggest that rather than reflecting arbitrary cultural standards, perceptions of facial attractiveness may convey biological information that is relevant to Darwinian 'fitness'. Here, computer graphic manipulations are used to test two evolutionary predictions. Firstly, that sexually dimorphic features (facial secondary sexual characteristics) are attractive in human male faces, as they may honestly advertise dominance and health through high androgen levels (expts. 1-5). Secondly, that facial similarity between the rater and the judged may predispose favourable attractiveness judgements due to the hypothesised benefits of assortative mating (expts. 6 & 7).

Experiments 1-5 indicate that women do not have clear preferences for exaggerated dimorphism in the form of 'masculinised' face shapes as predicted by indicator models of sexual selection. Unexpectedly, 'feminised' faces are preferred overall, and this effect is not due to artifacts in stimulus generation (expt. 2). Stereotypical personality judgements attributed to static faces appear to influence attractiveness judgements: masculine faces are perceived as possessing less desirable characteristics than 'feminine' male faces (expt.1). Female attractiveness judgements of male faces are mediated by menstrual cycle phase: when conception is likely, females prefer relatively more masculine face shapes (expt. 3). These cyclic preferences are, in turn, possibly mediated by personal situation (presence or absence of a partner (expt. 4) and the type of relationship sought (expt. 5))

A model linking likelihood of parental investment and facial masculinity preferred by females is suggested: when parental investment is sought (i.e. for a long-term relationship) facial shapes associated with relatively lower testosterone levels are preferred. Such faces may reliably indicate pro-social personality characteristics. When, however, likelihood of parental investment is low (in short-term relationships or possibly extra-pair copulations at times of high conception risk) relatively more masculine faces are preferred, in a fashion more consistent with 'good genes' hypotheses.

Experiment 6 showed a correlation between attractiveness and similarity but the effect can be explained by the attractiveness of average faces. Beyond this, there was a trend for individual subjects to rate opposite sex images with a similar face shape to their own face as more attractive than other subjects. Experiment 7 allowed subjects to interactively manipulate an opposite sex facial image along a continuum from a self-similar shape, through an average face shape, to a face with opposite characteristics. No significant preferences for self-similar or opposite characteristics were found. Preferences for average faces are stronger than preferences for self-similar faces.

1 Evolutionary Theories of Sexual Selection

1.1 Overview: Sex, fitness and evolution

Darwin's theory of natural selection, popularised by Spencer as the 'survival of the fittest', is challenged by readily observable characteristics in many species. Extravagant traits such as the peacock's tail seem to reduce an animal's chance of surviving so obviously that they are hard to explain by natural selection alone. Surely such a process would remove these traits from species, to generate organisms designed to survive and to reproduce, driven by efficiency. Yet ostentatious displays, primarily in males, are prevalent throughout the animal kingdom. Darwin's controversial theory of sexual selection was developed to explain such seemingly anomalous behaviour. Sexual selection theory explains that these traits, whilst not favoured by natural selection, give those that possess them a reproductive advantage. Darwin (1871) argued that male traits:

> "... serve only to give one male an advantage over another, for the less well-endowed males, if time would allow them, would succeed in pairing with the females; and they would in all other respects, judging from the structure of the female, be equally well adapted for their ordinary habits of life. In such cases sexual selection must have come into action, for the males have acquired their present structure, not from being better fitted to survive in the struggle for existence, but from having gained an advantage over other males... "

Sexual selection, then, is a subset of all selection pressures in the wild. Whilst some traits favoured by sexual selection may also be favoured by natural selection, others may not (Andersson, 1994). In this brief introductory chapter, several theories of sexual selection are introduced, and finally the possibility that sexual selection accounts for mate preferences in humans is assessed using a comparative approach.

1.1.1 A definition of fitness.

'Fitness', a key term in evolutionary discussions of behaviour, is best defined as reproductive success (viable offspring reared), rather than health or longevity per se. Survivability contributes to reproductive success only to the extent that it increases the chance of reproduction. Daly & Wilson (1983) suggest that fitness is, in fact, a three-part statement:

> "Variations in reproductive success exist among the members of populations; such variations are partly due to genetic differences and are therefore heritable; and these facts together explain the mechanism by which gene pools and phenotypes evolve... "

Reproductive success at the level of individual genes is the commonly accepted ultimate unit of selection within contemporary evolutionary biology (Dawkins, 1976). Genes that successfully reproduce themselves will replace genes that do not in the population. However, genes operate in conjunction with other genes at the level of the organism, and succeed or fail to reproduce together. Although intragenomic conflict may have serious evolutionary significance (Daly & Wilson, 1983; Andersson, 1994), and group selection arguments are again finding supporters (e.g. Wilson, 1997), the individual organism remains the most useful level of analysis for the evolutionary study of sexual behaviour.

1.1.2 Choosy females, competing males.

There is a basic asymmetry in mate choice in many species. Overwhelmingly, it appears that females choose, while males compete for sexual access: 'the males are almost always the wooers' (Darwin, 1871). The forms of this competition are many and varied, but whether the competition is based on scrambles, aggression or extravagant displays of secondary sexual characteristics, males compete for access to females in the vast majority of species.

This asymmetry is often thought to reflect anisogamy: the difference in size between male and female gametes. The origins of sexual reproduction are still disputed by biologists, and will not be discussed here. Once sexual reproduction exists however, disruptive selection has been proposed to account for anisogamy: intermediate sized gametes fail in competition against either very large or very small gametes (Parker, Baker & Smith, 1972). Females produce relatively few gametes, and those they do produce are large and energy rich. Conversely, males produce many smaller, more mobile sperm. In itself, this differential investment in gametes is enough to generate male competition for access to mates, as females who are choosy about their mate with will be more successful than less discriminatory females. The cost of failed mating is less for males with their relatively resource cheap, numerous gametes (Trivers, 1972).

Competition for a mate has been studied mainly in the context of polygynous species, in which male reproductive success is very variable: some males (often selected by females on the basis of secondary sexual traits) sire far more offspring than others, by monopolising females. Most mathematical models of sexual selection (discussed below) are based on assumptions of polygyny. Sexual selection can also occur in truly monogamous species where potential partners differ in parental ability or fecundity. In such species, where selecting a partner reduces a male's chances of achieving other mates, males and females are expected to demonstrate choice (Hamilton, 1990; Andersson, 1994). Assortative mating (see chapter 6) in combination with varying genetic quality can lead to fairly strong sexual selection for indicators of quality in species that form enduring pair-bonds. In addition, true monogamy is fairly rare in nature; most socially monogamous animals indulge in extra-pair copulations, undermining the apparently monogamous mating system and introducing increased variance in male reproductive success (Andersson, 1994). The next two sections introduce the two main theories of sexual selection - 'runaway' and 'good genes' models of sexual selection. Although initially seen as competitors, evidence has accumulated suggesting that both types of model may exist in reality, and are complementary rather than exclusive (Miller, in press).

1.2 Fisherian 'runaway' sexual selection

Runaway sexual selection occurs when a male trait and a female preference for that trait co-evolve in successive generations. In this process, first suggested by Fisher (1915), the origin of the preference for the trait rapidly becomes less important than the possession of the trait itself, leading to females choosing extraordinary male characteristics. Such a mechanism has been proposed to account for extravagant male plumage in birds (i.e. the peacock's tail) which is hard if not impossible to explain by natural selection alone.

Suppose that heritable variation in a male trait is linked to viability. For example, slightly increased tail length in birds could lead to greater survivability (Andersson, 1994). Females that preferred long tailed males would then have improved reproductive success: their male offspring would have longer than average tails, and hence a better chance of reproducing themselves, whilst their female offspring would inherit a preference for increased tail length. Alleles coding for both the trait and the preference become associated. This process may escalate, becoming a positive feedback loop in which increasing male tail length is accompanied by ever more choosy females who prefer long tails. Although slightly increased tail length may have originally been linked to increased survivability, this advantage may be lost as traits increase beyond an optimum size. Fisher proposed that under some circumstances, the reproductive advantages bestowed on males with exaggerated traits by sexual selection outweigh the selection against such traits by natural selection. Eventually, however, a very exaggerated trait leads to increased mortality in males and runaway sexual selection is brought to a halt by balancing natural selection.

Fisher's verbal conception of a runaway process has been modelled mathematically and genetically (see Andersson, 1994, for a review). O'Donald (1980) showed that an allele that reduces viability (e.g. an extra-long tail increasing predation risk) could increase in frequency in the population if the preference for it is widespread. Kirkpatrick (1982) and Seger, (1985) showed that traits that reduce viability can reach fixation in two loci mathematical models. Lande (1981) demonstrated that runaway selection could occur in a slightly more realistic polygenic model, although even this model relies on many assumptions such as mate choice carrying no cost, that are hard to justify (see Andersson, 1994). Despite being obviously simplified models of complex systems, mathematical and genetic models show that runaway sexual selection could evolve. The origin of the preference is not important to Fisherian processes: in the example above an increase in tail length was linked to improved viability. The initial preference could come from other sources, however, such as a sensory disposition evolved for another purpose (e.g. Ryan & Rand, 1990).

1.2.1 Empirical support for Fisherian selection

Empirical evidence in support of runaway sexual selection is, however, scant. Preferences for exaggerated male traits certainly exist, but are also consistent with other theories of sexual selection (see section 1.3 below). Several studies of guppies and sticklebacks show that a male trait and a female preference for that trait have co-evolved (Dugatkin & Godin, 1998; Andersson, 1994, for reviews), but perhaps the best evidence for a Fisherian process comes from Wilkinson & Reillo's (1994) study of stalk eyed flies. By selective breeding over generations of flies, they successfully bred one group of flies in which the males had long eye-stalks, and another group with short stalks. Females within each group preferred the stalk length of their own group: evidence that preference and trait had co-evolved.

1.3 Indicator Mechanisms of Sexual Selection

Indicator mechanisms of sexual selection are based on the assumption that a trait is preferred as it is associated with either phenotypic or genotypic quality. In contrast to Fisherian models of sexual selection, indicator models assume a trait is preferred due to a non-arbitrary fitness advantage associated with it. This fitness advantage is often considered to be heritable, and so indicator mechanisms are commonly known as 'good genes' models of sexual selection: "weak, ambiguous, unreliable and incidental cues of being non-dead and non-sick will become strong, clear, reliable... indicators of being vigorous and healthy" (Miller, in press). Energetically expensive conspicuous male traits are thought to represent an 'honest handicap' (e.g. Zahavi, 1975; section 1.3.2) and reflect the action of androgens, which may suppress immune system function (immunocompetence handicaps; Folstad & Karter, 1992; section 1.3.3). In species in which males contribute more than just genes to offspring, however, it is possible that male traits may advertise non-heritable phenotypic benefits, such as the possibility of improved parental ability through access to resources, discussed below.

1.3.1 Indicators of direct phenotypic benefits

Female choice for male parental ability could be an adaptive behaviour favoured by natural selection. In some bird species male traits such as colouration, song or display rate are influenced by the access a male has to food resources or to male age and experience. Experimental studies that indicate that these factors provide a direct advantage to females through improved paternal care are reviewed in Andersson, 1994. Exaggerated male traits probably reflect good phenotypic condition, which indicates a lack of parasitic infestation. Avoiding a parasitized male has an obvious direct advantage to females whether such parasite resistance is heritable or not (Gibson, 1990). Males in good condition may also prove superior in defending females from other males that may injure her. Mathematical models of female choice for non-heritable traits associated with direct phenotypic benefits suggest that such preferences could evolve (e.g. Heywood, 1989; Price et al, 1993; see Andersson 1994 for review).

1.3.2 Zahavi: honest handicaps

Zahavi's classic model of sexual selection considers that females have been selected to prefer exaggerated male traits as they are a handicap that only high quality males can bear successfully (Zahavi, 1975). Secondary sexual traits such as exotic plumage in birds or a stag's antlers are energetically expensive to produce: males must be in good condition to produce them. In addition, they often present a direct impediment to well-being: large traits compromise an individual's agility and ability to forage, and may also attract predators. Zahavi argues that by surviving such a handicap, males are demonstrating their quality. Poor quality males are incapable of diverting the resources necessary for the development of large secondary sexual traits, and so are unable to 'fake' the characteristics.

The strongest objection to Zahavi's handicap principle is that it relies on the heritability of fitness. Females who pick males on the basis of 'quality' (advertised by handicaps) will only benefit if their offspring also inherit this fitness advantage (Maynard-Smith, 1978). Theories of population genetics, however, suggest that this should not be the case (Fisher, 1930). If females prefer mating with males of high genetic quality, then soon offspring carrying the genes that lead to this fitness benefit will increase in number. Genes leading to fitness benefits in the population rapidly move to fixation: everybody has them. Hence, all differences in phenotypic condition come from environmental rather than genetic sources, and 'good genes' models of sexual selection become meaningless.

These criticisms of handicap theories, however, have not ended discussion of the possibility of indicator mechanisms. Some studies indicate that male ornamentation is related to offspring survivability (e.g. in Peacocks - Petrie, 1994; and in great tits Norris, 1993). Whilst many other empirical studies of the heritability of fitness appear inconclusive, Andersson (1994) points out that a very small amount of fitness heritability makes indicator mechanisms viable, and low heritability is difficult to demonstrate experimentally. Daly & Wilson (1983) note that such simplified models fail to take into account the possibility that combinations of genes may have non-additive fitness benefits, and that the overall level of heterozygosity in an individual (which is expected to vary between individuals) may influence fitness, and is heritable. Perhaps the greatest weakness of population genetics investigations of heritable fitness is that selection pressures in such models are assumed to be constant. Once this assumption is violated, the possibility that heritable fitness can be maintained in a population increases. One possible source of variation in selection pressures is parasites (e.g. Hamilton & Zuk, 1982). Extremely short parasite generation time sets up an 'arms-race' between host and parasite, leading to conditions where a successful genotype (i.e. parasite resistant) in one generation may not be so in the next (Thornhill & Gangestad, 1993). The role of parasites in 'good genes' models of sexual selection has attracted much attention and is discussed below.

1.3.3 Parasite driven sexual selection & the immunocompetence handicap

Hamilton & Zuk (1982) proposed that the varying selection pressure imposed by rapidly evolving parasites might allow fitness to have a heritable component. Their comparative study of North American bird species showed an associative relationship between blood parasites and plumage 'brightness'. Bird species with a heavy parasite load had more extravagant plumage, allowing males in good phenotypic condition to advertise their vigour through secondary sexual characteristics. In less parasitized species, Hamilton & Zuk suggest that there is less call for sexual selection of healthy males by female choice, and hence exaggerated ornaments are less common. Within several bird and fish species, male vigour and trait expression are reduced by parasitization (Zuk, 1992). Hence, female choice for exaggerated ornaments favours the least parasitized males. Although this may be directly advantageous to choosy females at the phenotypic level (see 1.3.1 above) some studies indicate that this resistance to illness and infestation may be heritable (e.g. in ring-necked pheasants: Hilgarth, 1990).

Researchers considering the role of sex hormones in sexual selection have elaborated the role of parasites. Males of many species, including humans, show higher levels of disease caused by parasites than females (Zuk, 1992), implying that testosterone, which suppresses acquired immune system responses, may negatively affect male health (Grossman, 1985)[1]. Testosterone, however, is the primary hormone responsible for the development of male secondary sexual traits, especially in mammals (Owens & Short, 1997). Hence, sexually dimorphic traits may be costly as they require the male to generate testosterone at the expense of the active immune system: male traits honestly advertise that an individual can successfully bear the load of the parasite burden even under a self-imposed stress. Folstad & Karter (1992) and Wedekind & Folstad (1994) present a feedback model in which testosterone production is a plastic response to parasite load: when possible, individuals reallocate resources to develop ornamentation to attract mates. Wedekind (1992) shows that ornamentation reveals very precise information about parasite load in roach, with different parasites preventing the expression of different male traits.

The evidence for immuno-suppressive properties of testosterone is limited, although there is research that supports the mechanisms of parasite mediated sexual selection when testosterone is administered at pharmacological doses in laboratory conditions (Hillgarth & Wingfield, 1997). Some empirical studies show physiological levels of androgens suppressing the immune system in birds, and that steroid levels drop in response to parasite infection in mice and rats (see Hillgarth & Wingfield, 1997, for a recent review of testosterone and immuno-supression in vertebrates).

---

[1] Males may also suffer increased parisitization due to lifestyle differences rather than direct hormonal causes. Life history studies indicate men live a 'riskier' lifestyle than women, especially in early adulthood (Daly & Wildon, 1983).

1.3.4 Developmental stability, fluctuating asymmetry and fitness.

Indicator mechanisms of sexual selection obviously require indicators to advertise 'good genes'. One such indicator is proposed to be developmental stability: the ability of an individual to develop successfully in the face of environmental pressures, such as illness. Bilateral body traits (e.g. limbs, digits) in the majority of species are symmetric when averaged across individuals. However, as the same genes code for both left and right side development, deviations away from symmetry provide an index of the success of an organism's growth. Developmental instabilities (manifested in increased fluctuating asymmetry, FA) are known to be caused by various inclement environmental (e.g. climate, malnutrition, parasitization) and genetic (e.g. inbreeding: see section 6.2.2) factors (Møller , 1997; Parsons, 1990).

Individuals of many species clearly assess developmental stability (either directly, through the amount of FA, or through some co-variate of symmetry) in potential partners: symmetrical individuals are preferred in many species (Møller , 1993; Manning & Chamberlain, 1993; see Møller & Thornhill, 1998 for a review). In addition, other fitness related characteristics (growth rate, fecundity and survivability) are positively associated with symmetry across species and taxa (Møller , 1997). Critically for 'good genes' theories of sexual selection, meta-analysis indicates that developmental stability is heritable (Møller & Thornhill 1997a, 1997b; but see responses of other authors in J. Evol. Biol., 1997, vol.10, pg. 17-67 for comments on this analysis).

Developmental stability is positively associated with trait size: within and across bird species larger ornaments (e.g. tails) tend to be more symmetrical than smaller ornaments (Møller , 1992; Møller & Hoglund, 1991; Andersson, 1994). This initially seems to be counter-intuitive, as the high androgen/oestrogen ratios implicated in such ornamentation should impose a developmental stress on their owners, resulting in increasing asymmetry with ornament size. However, the finding can be explained by a resource allocation model of ornament growth (section 1.3.3 above) - only individuals that are capable of bearing the immunocompetence handicap do so. In addition, this finding seems to be more consistent with a 'good genes' model of sexual selection than an arbitrary runaway process (Møller & Hoglund, 1991; Andersson, 1994).

1.4 Sexual selection in humans

1.4.1 Human mating system & sexual dimorphism: comparative approaches

The mating system adopted by a species can be used to assess the likely strength of sexual selection (Andersson, 1994). Whilst some degree of sexual selection is predicted in monogamous species (section 1.1.2 above) stronger male competition is expected in polygynous species, where one male secures a large proportion of copulations. In polygynous species sexual dimorphism tends to increase as a result of more intense sexual selection, either through either male-male competition or female choice.

Humans demonstrate a wide variety of mating systems, from polyandry to polygyny. Anthropological studies indicate that most human societies in Murdock's (1967) ethnographic atlas allow polygyny (708/849), but within these cultures the vast majority of marriages are monogamous, with only very high status men marrying multiple wives[2]. Many researchers are quick to point out that marriage system, however, may differ somewhat from actual mating system in many monogamous societies. Although reliable figures are hard to come by (Macintyre & Sooman, 1991), some degree of extra-pair paternity (where an infant's biological father is not the 'social' father) seems to be found all in cultures that have been studied (e.g. Kurland, 1979; Baker & Bellis, 1995; Buss, 1999). Cryptic ovulation in human females may be an adaptation to introduce paternity uncertainty in human offspring – a theory analysed in depth in Chapter 5, below.

Human males and human females are moderately sexually dimorphic. Modern males are 5-12% taller than females (Alexander et al, 1979) and weigh approximately 20% more. Alexander et al (1979) review sexual dimorphism data from a variety of mammal species (including ungulates, pinnipeds & primates). A clear relationship between both mean and maximum harem size and the amount of sexual dimorphism is found: species in which females are smaller than males tend to be polygynous. Alexander et al conclude from this comparative analysis that human sexual dimorphism in body size falls within the range of a 'mildly polygynous species'.[3]

---

[2] Marriage system appears to be related to pathogen stress in human societies (Low, 1990). In societies with high levels of pathogen prevalence, polygyny is more common. This finding has obvious implications for immuno-competence handicap theories of sexual selection in humans: when the threat of illness increases, the value of 'good genes' would increase, and the 'costs' of polygyny (presumably lower paternal investment in offspring) may be out weighed by the benefits of genetic quality.

[3] Alexander et al (1979) also report differences in levels of sexual dimorphism between separate human populations, although the interpretation of such data is problematic. Human populations in marginal environments tend to be monogamous. Alexander et al consider monogamy in these situations to have been ecologically, rather than socially, imposed. Their analysis indicates that in such groups, degrees of sexual dimorphism are reduced.

A comparative analysis of testes weight in primates suggests that human testes are an appropriate size for uni-male polygyny or pair bonding mating system: testes tend to be larger in species with polygynous-promiscuous multi-male mating systems (Harcourt et al, 1981). In addition, men reach sexual maturity at a later age than women. Comparison with other species indicates that such sexual bimaturism occurs when younger males are unable to compete with older males for sexual access – an indicator of polygyny (Daly & Wilson, 1983; Andersson, 1994).

The fossil record has been used to infer the mating system of ancestral hominids. For example, *Australopithecus afarensis* is estimated to have been far more dimorphic than *Homo sapiens*, with the male-female difference approximately equal to the differences found in gorillas today (Lockward et al, 1996), suggesting a uni- or multimale/multifemale mating system. A parsimonious phylogenetic analysis by Sillen Tullberg & Møller (1993) suggests that the species ancestral to *Homo* were probably harem holders, and that multimale mating systems evolved independently in chimpanzees. In summary, comparative analyses of humans with other species and fossil evidence suggest that levels of initially high levels of polygyny within an ancestral mating system have been decreasing, leading to decreasing levels of sexual dimorphism from ancestral hominids to modern humans.

1.4.2 Hunting and the evolution of the human mating system

Humans, relative to many other animals, invest heavily in few offspring rather than producing many progeny with little investment in each one (Humans can be considered to be extreme K-strategists, in comparison r-strategist species with short life-spans and large numbers of offspring; Pianka, 1970)[4]. The combination of changing anatomy of the pelvis to accommodate bipedalism, and increasing adult brain size lead to increasing immaturity of the hominid neonate, as large brained infants would not fit through the shrinking birth canal (Daly & Wilson, 1983). An immature infant needs extensive parental care, and modern humans of both sexes provide such an investment in their offspring. Pair-bonding is suggested to have evolved out of a polygynous mating system fairly early in hominid evolution, possibly due to the benefits of bi-parental investment in relatively helpless human neonates (Daly & Wilson, 1983).

Brain tissue is energetically expensive, and the cost is born primarily by the mother during pregnancy and weaning (Foley & Lee, 1991). Pair-bonding, often formalised as marriage, is "not in essence a sexually based behavioural association between a male and a female, but rather an economic and child-rearing partnership" (Symons, 1979). Many theorists assume that hunting and meat provisioning by males in addition to female foraging allowed the rapid encephalization of ancestral hominids (e.g. Symons, 1979; Buss, 1992).

Miller (1997), however, cites Foley & Lee's (1991) data concerning the energetic demands of human, as opposed to chimp neonates. Their data suggests that the

---

[4] MacArthur's *K*- and *r*- strategies (1962) refer to life history patterns that develop in response to environmental pressures. The term *r* is a measure of the rate of increase in the size of a population. *K* refers to the carrying capacity of a habitat: the maximum population size it can sustain. *r*-selected species attempt to maximize *r* by maturing early and producing many offspring as the evolutionary environment has repeatedly sustained rapid population growth. *K*-selected species have a population size at or near the carrying capacity of the habitat, and whose life history and breeding system reflects survival and reproduction when the population is around *K*. Daly & Wilson (1980) review criticism of *r*-/*K*- selection theories.

additional energetic demands of human over chimp growth are to some extent born by the evolution of slower growth rates: "human infants do not grow using more energy per month than other ape infants, as paternal provisioning would have made possible – they simply grow for a longer time" (Miller, 1997). Miller claims that this allowed ancestral females to choose males on the basis of heritable genetic quality ('good genes' sexual selection) rather than on the basis of likely paternal provisioning.

Miller's theory, however, disagrees somewhat with Foley & Lee's own interpretation of their data. Foley & Lee estimate that maternal energetic effort has increased by 17% over the course of hominid evolution. Whilst they agree that slower infant development may account for some of this cost, human infants still require approximately 8.7% more energy per day than chimpanzee infants up to the age of 18 months. Improved hominid foraging techniques may make up some of this shortfall, but Foley & Lee conclude that 'higher quality foods in substantial quantities in the diet', such as male provisioning of meat, were still probably necessary over the course of human evolution (see section 1.4.3 immediately below for an alternative source of an improved diet). Buss (1991) points out that even if male provisioning were not necessary for the raising of offspring, it would certainly have been advantageous, and as such would be favoured by selection. As mentioned above, decreases in sexual dimorphism also implicate a shift toward more pair-bonding over hominid evolution.

1.4.3 The Grandmother hypothesis

The common belief that male hominid hunting of large game was a catalytic event in hominid evolution may, however, be flawed. The provisioning of offspring with meat obtained by large game hunting has been hypothesised to have led to nuclear families and 'modern' sexual division of labour. This theory originates from archaeological evidence of stone tools (fashioned by *Homo erectus*) associated with the remains of large animals, and evidence of improved diet from the reduced size of the gut in *H. Erectus*. From this evidence, researchers developed theories of central place foraging in which male hunters killed and butchered large game and transported it to share with family members (e.g. Isaac, 1978).

Re-evaluation of the archaeological evidence and extensive studies of extant hunter-gatherer groups dispute this widely held belief (O'Connell et al, 1999). Studies of modern hunter-gatherers (especially the Hadza) indicate that large game hunting is too variable a strategy to reliably feed offspring, and kills, once made, are indefensible from other group members. If male provisioning of offspring of meat was an important factor in hominid familial organisation, small game hunting would probably have been a more efficient strategy (O'Connell et al 1999; Hawkes et al, 1991). Instead of big game hunting, O'Connell et al propose that the benefits of foraging for new food sources, specifically tubers, led to a selection pressure that increased the post-reproductive life-span of females. In this model, the advantages proffered by an active foraging grandmother providing a reliable food source for her daughter's children are so great that senescence is strongly selected against. Evidence from the Hadza shows that grandmothers well into their sixties provide significant contributions to their kin's diets through foraging. It is possibly this life history change, rather than a new food source in the form of meat, that 'catalysed' hominid evolution. The grandmother hypothesis, however, fails to explain the decreasing sexual dimorphism in the hominid line, and the assumed change in mating system that accompanied it.

1.4.4 Human mate choice: advertising females and choosy males?

Whatever evolutionary events led to current levels of sexual dimorphism or the life history changes that increased the length of time that humans reach maturity, both factors predict that some degree of mate choice by both men and women should occur: significant investment by both parents makes indiscriminate mating costly. Physical beauty seems intuitively to be an important element in partner choice. As male reproductive success varies more than female reproductive success in humans (either through polygynous mating system or extra-pair sex), theories of sexual selection and cross-species comparisons indicate that males should advertise physical characteristics more than women.

Cross-cultural studies of mate preferences, however, indicate that physical appearance is more important to males than females in a partner (Buss, 1989). Observation suggests that women, rather than men, are the sex that advertise through extensive use of makeup and clothing: "Faced with these facts, a biologist would be forced to suspect that he was looking at a society in which females compete for males, rather than vice-versa" (Dawkins, 1976).

Despite Dawkin's observation of 'female display' above, the idea that males are the choosier human sex is probably misguided. Symons (1979) argues that female reproductive value is more easily estimated from physical appearance than male reproductive value, and hence females advertise physical characteristics more than men. Young, healthy women are most likely to have a long reproductive career ahead of them.[5] Cues to youth and health are indeed found attractive in female faces and bodies (see next chapter for a review of the facial attractiveness literature; also Symons, 1979, 1994; Singh 1993a, 1993b, 1994, 1995; Tovee et al, 1998).

Male ability to reproduce declines less rapidly with age than females, and yet access to resources may increase with age through status. There is a double standard of aging: women's attractiveness decreases with age more rapidly than men's (Deutsch et al, 1986). Older men may have advantages over younger men in and have access to resources though male-male competition. Females seem to find cues to status more important than males in potential partners. Pawlowski & Dunbar (1999) analysed the age preferences specified in personal advertisements placed by men and women in a British newspaper. From this analysis, Pawlowski & Dunbar estimated that female 'market value' peaks in the 20s, and males reach their peak market value in their late thirties. Evolutionary psychologists attribute this to the greater capabilities of high status males to invest in their offspring (Kenrick & Keeve, 1992; Symons, 1979; Buss & Barnes, 1986; Buss, 1989; 1999; Townsend & Wasserman, 1998). Women, then, appear to be attracted to potential direct phenotypic benefits in males.

Despite relying less on physical appearance for cues to attractiveness than males, females still appear to be the 'choosier' in terms of actual sexual activity, as predicted by parental investment theory (Trivers, 1972; Symons, 1979; Townsend et al, 1995). Whilst it is clear that neither males nor females are indiscriminate in their choice of long-term

---

[5] Physical attractiveness in a partner is especially important in human societies that are subject to a higher parasite load (Gangestad & Buss, 1993). Physical attractiveness (advertisement of genetic quality within good genes theories) is undoubtedly an potential influence on both male and female mate choice.

mate, the differences in minimum parental investment between males (a single copulation) and females (pregnancy) ensure that the basic asymmetry between males and females found in most species is maintained in humans (Buss, 1989; Quellar, 1993; Townsend et al, 1995). Female American college students report fewer desired lifetime sexual partners and are less likely to report sexual activity without emotional attachment than males (Buss & Schmitt, 1993; Townsend et al, 1995).

1.4.5 Sexual selection and human faces

Human anatomy suggests a mildly polygynous mating system. Ethnographic and behavioural data support this suggestion: even in societies with apparently monogamous marriage systems, some degree of extra-pair paternity exists. Observation and introspection suggest that human faces, which are sexually dimorphic, are an extremely important influence on mate choice for both men and women. Faces may also provide biological cues to mate value through the advertisement of health, dominance or other desirable characteristics. The next chapter discusses theories of sexual selection with respect to human facial attractiveness judgements.

Sophisticated mate choice decisions will certainly balance cues to both attractiveness ('good genes'), status and willingness to invest in offspring ('direct benefits'); ecological pressures may influence these decisions.

2 Sexual Selection and Human Facial Attractiveness

2.1 Influence of facial characteristics on behaviour.

Facial appearance influences many aspects of our social and sexual lives. The face conveys so much social information that its importance in everyday life cannot be overestimated. Attractive faces seem subject to pervasive 'halo effects': positive characteristics attributed to attractive individuals give apparently unwarranted advantages to 'beautiful people' (e.g. Feingold, 1992).

Whilst both males and females claim in self-report that physical attractiveness is not of primary importance when choosing a partner (Buss, 1989), the single best predictor of satisfaction with a 'blind date' is facial attractiveness for both men and women (Walster et al, 1966). Given the apparent importance of the face in mate choice decisions, and the centrality of mate choice theories to evolutionary approaches to behaviour, theories of sexual selection outlined above may have acted on human facial characteristics. This chapter reviews studies of facial attractiveness indicating that the face may contain valuable biological cues to mate quality.

2.2 Cross-cultural and developmental studies of facial attractiveness

The first goal of researchers proposing a biological propensity to find certain faces attractive is to demonstrate that the learning of arbitrary cultural values has little influence on judgements of attractiveness - that attractive faces are considered so universally. Researchers have demonstrated high cross-cultural agreement in ratings of attractiveness of faces of many ethnicities. Cunningham et al, (1995), showed that

Hispanic, Asian, Black and White subjects judgements' of attractiveness of photographed individuals correlated, with coefficients varying between 0.91 and 0.94. These correlations were irrespective of the ethnicity of the photographs judged. Perrett et al (1994) found consistency across Japanese and Caucasian ratings of both Japanese and Caucasian stimuli. Jones (1995) showed that males from five populations (Brazilians, Americans, Russians, Paraguayan and Venezualan Indians) were all attracted to the same features in female faces. Bernstein et al (1982) found 'essentially no differences' between cross- and within- ethnicity attractiveness judgements using Black, Chinese and White subjects. Even Darwin (1871) acknowledged the possibility of the universality of facial attractiveness across cultures when he noted that explorers had remarked that indigenous peoples around the world had similar standards of beauty.

A second line of evidence in favour of a biological rather than an arbitrary cultural basis of physical attractiveness judgements comes from studies of infant preferences for face types. Langlois et al (1987) showed pairs of female faces to infants between 2-3 months and 6-8 months old. The attractiveness of the faces had been judged by adults prior to the images presentation to the infants. Infants preferred to look at the more attractive face of the pair, indicating that even at two months of age, adult-like preferences are demonstrated. A further study replicated and extended the finding to show that infants also expressed adult-like preferences for Caucasian male faces, Black female faces, and even the faces of other infants (Langlois et al 1991). Slater et al (1998) demonstrated that newborn infants prefer attractive faces.

Infants have been demonstrated to form cognitive prototypes extremely rapidly: Exposure to exemplar faces in early infancy may allow infants to learn facial configurations (de Haan et al, in review). So, although infants demonstrate adult like preferences very soon after birth (e.g. Slater et al, 1998), this doesn't necessarily imply that children are born with 'innate' knowledge of attractiveness. Nonetheless, the learning process occurs rapidly, without any linguistic component and at a very young age. Such learning could be considered to be effectively 'culture free' and hence possibly more analogous to an imprinting process than the learning of cultural norms.

2.3 Sex differences in facial growth

Sexual selection is often implicated in sexual dimorphism (Chapter 1 above). Before discussing studies of human facial attractiveness, it will be useful to be aware of the sexual dimorphisms found in human faces. Male-female dimorphism in human faces radically increases at puberty under the influence of sex steroids, particularly testosterone and oestrogen. Levels of both hormones increase in both pubescent boys and girls, although the ratio of androgens to oestrogens is dependent on sex. High testosterone levels causes forward growth of the brow ridges, and an increase in the size of the bones of the jaw, lower face and cheekbones (Thornhill & Gangestad, 1996; in prep; Gage et al, in prep). As masculinity in mammals is testosterone dependent, whereas femininity represents the more neutral developmental state (Haaq & Donohoe, 1998, Owens & Short, 1995), male faces grow more than female faces, which remain relatively childlike. High oestrogen apparently prevents facial bone growth in pubertal females, and leads to fat deposition in the lips and cheek area (Thornhill & Gangestad, in prep).

2.4 'Good Genes' theories and human faces

Despite increasing professional and public interest in the field, systematic studies of facial attractiveness have not yet reached consensus on the characteristics that make human faces attractive. The three main hypotheses are that (a) symmetrical faces (b) average faces, or (c) faces with exaggerated secondary sexual characteristics are found attractive. Each of these hypotheses has can be justified by 'good genes' theories, and each has received some empirical support, briefly reviewed below.

All three of these theories rely on facial characteristics advertising biological quality honestly: symmetry, averageness or exaggerated sex-typical traits are thought to advertise characteristics linked to reproductive potential, and hence preferences for such characteristics are favoured by selection. Empirical proof that facial attractiveness is associated with characteristics that advertise health is, however, scant at this time. One large-scale study of facial attractiveness found no link between facial attractiveness and health (Kalick et al, 1998) in a modern American population[6].

2.4.1 Symmetry

Symmetry appears to be associated with sexual selection and reproductive success of humans (Gangestad and Thornhill, 1997, Møller et al, 1995; Singh 1995; Thornhill et al 1995). Measures of symmetry in the human body and face correlate with attractiveness (Gangestad et al. 1994; Grammer and Thornhill 1994). Such correlations may, however, be due to other factors co-varying with symmetry. For example, sex hormones may influence the symmetry of growth (Thornhill and Gangestad 1993) and chin shape and size, which independently affects attractiveness (Perrett et al 1994; Scheib et al, 1999).

---

[6] Daly & Wilson (1999) note that selection pressures in 20th century San Francisco may not reflect the environment in which the proposed psychological mechanisms for detecting facial attractiveness evolved. A poor correlation between attractiveness and health in a modern population does not necessarily indicate that in harsher environments such a link would not exist.

2.4.1.1 *Studies of naturally occurring facial asymmetries*

There is no standardised technique for measuring the absolute symmetry of a human face. The difficulty involved in an accurate measurement of symmetry may be a factor that contributes to the relatively small amount of research that has been conducted into the symmetry of natural faces. Grammer & Thornhill (1994) estimated the asymmetry in male and female faces by bisecting lines that delineated left and right positions of bilateral features, marked on each face in a small sample of 16 men and 16 women. Horizontal symmetry measured in this way correlated with attractiveness judgements of both male and female faces.

Using a similar measurement technique, but with the inclusion of a measure of vertical as well as horizontal symmetry, Scheib et al (in press) also found that symmetry and rated attractiveness correlated for a sample of 40 male faces. Fascinatingly, the relationship between symmetry and facial attractiveness was still observed when only the left or right half of each face was presented (i.e. without cues to symmetry). This suggests that some co-variate of symmetry that can be ascertained from half faces may influence attractiveness judgements. Scheib et al found that size of the lower face and cheekbone prominence was related to symmetry and both half- and full-face attractiveness judgements. However, it should be noted that some cues to symmetry may still be present in half-faces. The mid-line used to generate the half-face stimuli bisected the middle of the nose tip and the 'v' of the upper lip. In faces with low symmetry, other points on this mid-line may reveal more (or less) than half of centrally placed features (e.g. bottom of chin; bridge of nose). Hence, some cues to symmetry may still remain in half-faces.

Mealy et al (1999) studied the co-variation of symmetry and attractiveness in monozygotic twin pairs. Such twins are genetically but not developmentally identical, and hence manifest differing levels of facial symmetry when adult. Symmetry of each twin pair was assessed by presenting left-left (L-L) and right-right (R-R) chimaeric image pairs of each of the two twins to subjects, and asking them to rate the similarity of each pair of images (Twin 1 L_1-L_1 and R_1-R_1; Twin 2 L_2-L_2 and R_2-R_2). The most similar pair of images (i.e. L_1-L_1 and R_1-R_1 or L_2-L_2 and R_2-R_2) was assumed to belong to the more symmetric of the two twins. Secondly, the natural images of the twins were presented to separate subjects, who were asked to rate the more attractive of each twin pair using a rating task. The results indicated a significant correlation between symmetry and attractiveness – the more symmetrical of the each twin pair was generally preferred, and bigger differences in estimated symmetry led to larger attractiveness differentials within twin pairs.

2.4.1.2 Studies using computer graphic symmetry manipulations

Given the theoretical benefits ascribed to symmetry (e.g. Møller & Thornhill, 1998; section 1.3.4 above), and the results of studies of naturally occurring asymmetry in faces, it is surprising that several studies directly manipulating human facial images have found that asymmetry is generally preferred to symmetry (Kowner, 1996; Langlois et al, 1994; Samuels et al, 1994; Swaddle & Cuthill, 1995). Most of these studies have created symmetric face images by aligning one half face with its mirror reflection (Kowner 1996; Langlois et al, 1994; Samuels et al, 1994). These techniques may induce additional stimulus differences unrelated to symmetry. The mirror reflecting technique can introduce abnormal feature shapes. For example, a mouth of normal width displaced to the right of the midline (Figure 1a) will assume atypical widths in left mirrored (Figure 1c) and right mirrored (Figure 1d) chimaeric face images. Asymmetry may have been preferred in a further study (Swaddle & Cuthill, 1995) because asymmetric original faces were compared to symmetric images with different skin textures (induced by combining original and mirror images). Whilst averaging textures over large number of faces results in an even skin texture, combining a face with it's mirror image may actually increase the number of apparent blemishes. For example, a face with a dark spot on the left cheek combined with its mirror image would generate a face with two slightly paler symmetrical spots on the left and right cheeks.

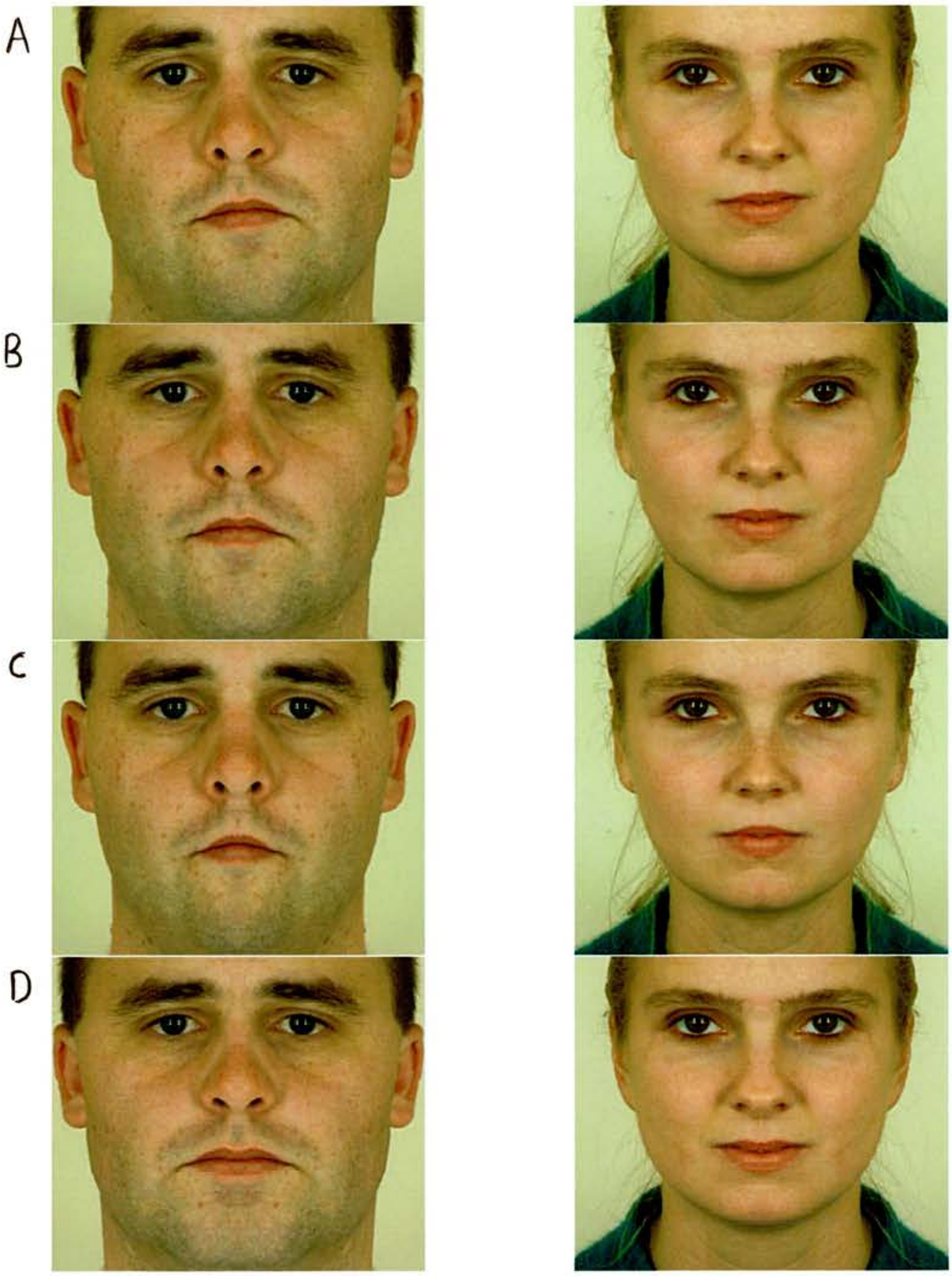


Figure 1. Symmetry manipulation for facial images with natural skin textures. Real face images with normal and symmetric shapes. (a, b) Note that asymmetries in pigmentation and shadows present in the original faces (a) remain in the more symmetrically shaped versions (b). (c, d) Images made with techniques employed in previous studies of facial symmetry. (c) Chimaeric faces made by combining the left sides of the original faces with their mirror reflections (c) and similarly for the right sides of the faces (d) illustrate the shape abnormalities that this technique induces. From Perrett et al (in press)

Alternatively, humans, unlike many animals, may prefer some degree of asymmetry in faces because this makes them look more distinctive or expressive. In the population there are tendencies for directional asymmetries (consistent differences between left and right sides that occur throughout the population). The most prominent directional asymmetries are, however, transient in nature and occur during speech and emotional expressions. During speech, most people (76%) make movements of greater amplitude on the right side of their mouth (the left side on an image, Graves et al 1982; Graves & Landis 1990; Wolf & Goodale 1987). These asymmetries are probably related to the greater strength of contralateral neural connections between the left (linguistic implicated) hemisphere and the right side of the face[7]. These naturally occurring asymmetries could contribute to attractiveness and may explain the preferences that have been reported for asymmetric human faces.

Despite the possibility that some degree of asymmetry may be attractive in human faces, several studies have demonstrated that symmetry predisposes favourable facial attractiveness judgements, reconciling the results of computer graphic studies with studies of naturally occurring asymmetrry. Perrett et al (in press) present three experiments using novel techniques for manipulating symmetry: these experiments are described in some detail below.

In experiment 1 symmetry in face shape was improved without changing the symmetry of face textures; natural asymmetries in skin pigmentation were present in

---

[7] During emotional expression, the left side of the face appears more expressive than the right (e.g. Skinner & Mullen 1991). Most researchers (e.g. Borod et al 1981; Borod & Koff 1983; Wylie & Goodale 1988) have found that spontaneous expressions are more asymmetrical than posed expressions though the findings are debated (Skinner & Mullen 1991).

both original and more symmetric versions of the same face (figure 1: section 7.4 below). Subjects' responses to pair-wise presentation of these two versions of each face indicated a clear preference for the symmetrically remapped stimuli (see section 7.4 for details of techniques). Experiment 2 used stimuli with average texture information generated from a set of faces. This average texture was rendered into both original and symmetrically remapped versions of a set of individual faces, giving perfect symmetry in the remapped version (figure 2; section 7.4). Pair-wise presentation showed a preference for perfectly symmetrical face stimuli. Experiment 3 used a rating task rather than a forced choice paradigm (stimuli were presented one at a time rather than in pairs), replicating the preference for symmetry shown in experiments 1 & 2.

The results of experiment 3 closely parallel those of experiment 2 despite the different methods of assessment. In experiment 1, symmetry manipulations were rather subtle; perhaps because of this subtlety of the manipulation, 75% of subjects were unaware that symmetry had been manipulated. Experiment 1 goes some way to establishing the ecological validity of computer graphic research into facial symmetry

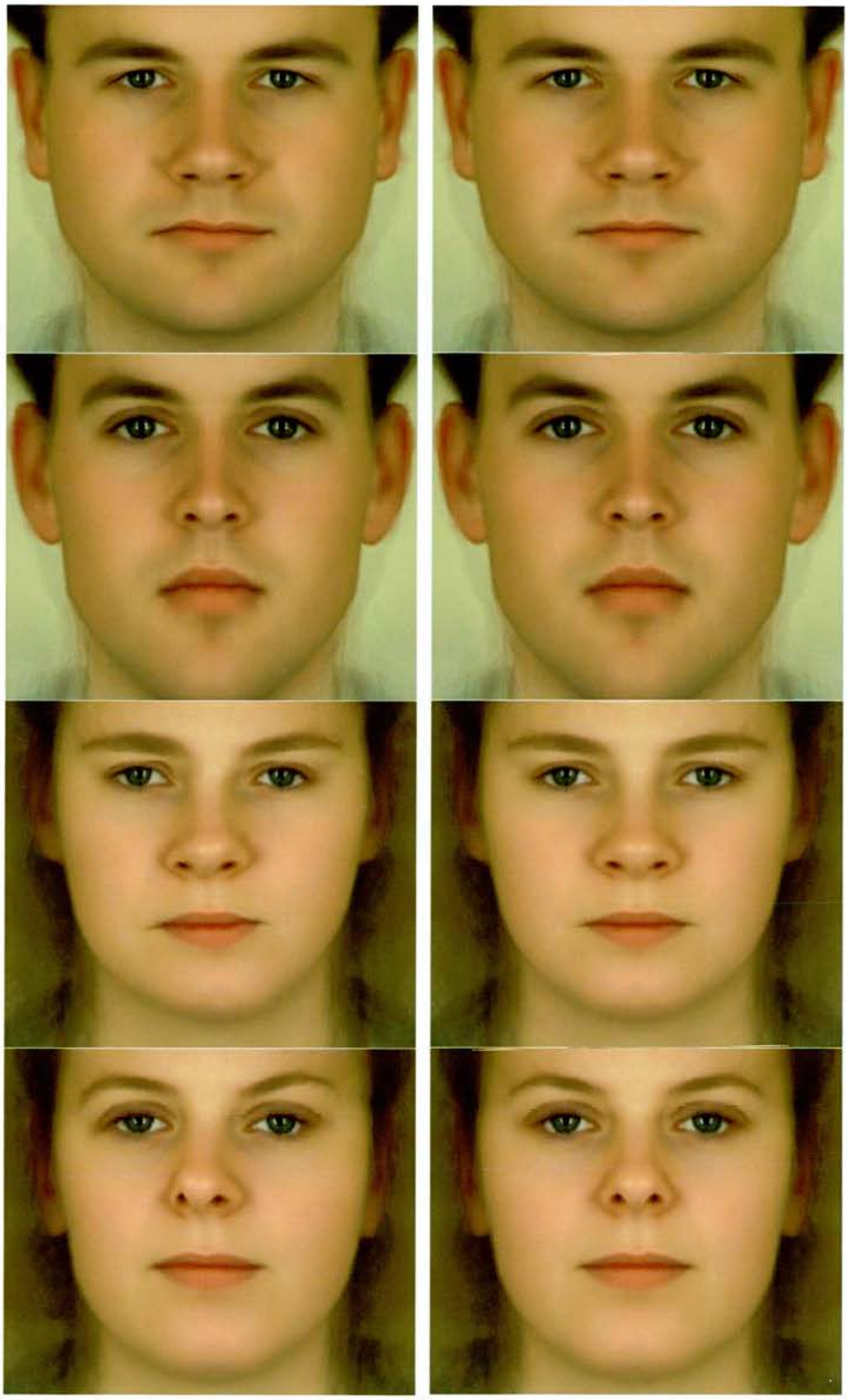


Figure 2: Asymmetrical and symmetrical face images with average facial textures. Normal (left) and symmetrical (right) face shapes with constant, average facial textures. Average texture information was produced by warping and blending 15 faces and their mirror images into the original, asymmetrical face shapes (left) and the symmetrical shapes (right). From Perrett et al (in press)

because it demonstrates a preference for more symmetrical facial images with entirely natural skin pigmentation and texture.

The use of such natural skin textures in Experiment 1 did not allow the face images to be made completely symmetric. Therefore Experiment 1 leaves open the possibility that complete symmetry might be less attractive than a slight degree of asymmetry. With synthetic skin textures, Experiments 2 and 3 show that perfectly symmetrical faces are preferred to faces with normal levels of shape asymmetry.

Rhodes et al (1998) used a technique similar to Swaddle & Cuthill (1995), but avoided stimuli artefacts such as double blemishes (see above) by retouching the manipulated images: double blemishes were removed using a computer graphics package. Using this improved technique, they too found a preference for symmetry over asymmetry. The result of Rhodes et al complement Perrett et al (in press) in that both demonstrate that asymmetry in faces reduces attractiveness. Further support for a preference for symmetry in human faces comes from Hume & Montgomerie (1999) who replicated Rhodes et al's technique with similar results.

There are progressive changes in the shape of faces, particularly in soft tissues, throughout life (e.g. Burt & Perrett, 1995). Asymmetries in face shape due to differential growth and ageing will therefore be more prominent in older faces. Kowner (1996) demonstrated that asymmetry itself may be a perceptual cue to age since asymmetric faces were perceived as older than symmetric faces.

It is apparent that symmetry is one factor contributing to the spectrum of influences on facial attractiveness. It is also apparent, however, that symmetrical faces may posses other properties that are associated with attractiveness, such as exaggerated sex-typical characteristics, and that symmetry may not be a primary cue to attractiveness judgements (Scheib et al, in press).

2.4.2 Averageness

The most influential theory of facial attractiveness is that average features, not exaggerated characteristics, are optimal, as such faces are thought to indicate high levels of heterozygosity in their owners (Symons, 1979; Thornhill & Gangestad, 1993). Genes in an individual lead to the development of proteins that form the environment for parasites. These pathogens are generally best adapted to proteins that are common in the host population. Rare alleles in a host may result in rare proteins that pathogens are poorly adapted to. Therefore, individuals with rare alleles may have a greater resistance to parasitic illnesses. A preference for average faces may be adaptive due to selection pressures operating on extreme genotypes that are more likely to be homozygous for deleterious alleles (Symons, 1979) and less likely to have alleles to which pathogens are poorly adapted (Thornhill & Gangestad, 1993). A preference for averageness is compatible with both cognitive theories of prototyping (see Langlois & Roggman, 1990) and work in theoretical biology that suggests a preference for 'average' phenotypes would rapidly replace random mating (Koeslag, 1994a; Koeslag, 1994b). However, the averageness hypothesis has received only mixed empirical support.

Modern computer graphic techniques develop a method first suggested by Francis Galton (1878) to allow composite 'average' faces to be constructed from much larger sets of individual photographs. Langlois & Roggman (1990) found that composite (average) faces were rated as more attractive than the individual faces that formed them. A preference for average faces over extremely distinctive faces has been demonstrated using line drawings (Rhodes and Tremewan, 1996). This preference may explain the popular belief that people are attracted to opposite sex individuals that resemble themselves. Faces very far from average receive low attractiveness ratings, these same faces resemble the faces of subjects less than more average faces (Penton-Voak et al, 1999; Chapter 6 below).

Alley & Cunningham (1991) proposed that average sized features would be attractive, but not optimal, and briefly review studies using methodologies not reliant on digital averaging that failed to support the averageness hypothesis. Perrett et al (1994) demonstrated that, whilst average composite male and female faces are indeed attractive, more attractive composites can be generated by using a subset of the 15 most attractive faces from a set of 60 individuals group. A counter argument to this finding suggests that the 15 most attractive individuals in the group may be, in fact, the 15 most 'average' faces in the group, and hence a composite of them represents the population average better than the full 60 face group. This argument however, cannot explain that a caricature of the 15 face average (exaggerating the differences between the 60 and 15 face average) creates a yet more attractive face (Perrett et al, 1994).

2.4.3 Averageness and symmetry: the impossibility of dissociation.

Average composite faces tend to have smooth skin and be symmetric; these factors, rather than averageness per se, may lead to the high attractiveness attributed to average faces (Alley & Cunningham, 1991; Benson & Perrett, 1992). Symmetry appears to contribute to attractiveness judgements (Perrett et al, in press; section 2.4.1 above).

Dissociating symmetry and averageness is problematic. As no reliable directional asymmetries have been found in human faces at rest, average faces are by definition symmetric. However, average size and symmetry of real faces may not be simply related. Studies of bird species indicate that larger than average secondary sexual characteristics (such as swallows tails) are more symmetric than smaller traits (section 1.3.4 above). In women, breast asymmetry is negatively related to breast size (large breasts are relatively more symmetric than small breasts; Manning et al, 1997) and in men facial symmetry appears to be related to jaw and cheekbone size (Scheib et al, in press). So, although in composite images symmetry increases as facial averageness increases, this may not reflect patterns of symmetry in real faces. Symmetry, rather than averageness, may drive the attractiveness judgements of composite faces. Pollard et al's 1999 study used facial-metric measurements of feature size rather than composites to calculate averageness, and found that faces with features close to the mean size of the population studied were rated as average, rather than high, in attractiveness.

2.4.4 Secondary sexual characteristics and sexual dimorphism

Symons (1994) proposes that averageness may be a 'default' position that defines facial attractiveness unless another psychological mechanism is operating. A preference for secondary sexual characteristics may be one such alternative mechanism operating in females' judgements of male facial attractiveness, and male's preferences for female faces.

Considerable evidence suggests that extremely feminine female faces are considered attractive. A wide variety of techniques ranging from measurement of facial photographs of women (Cunningham et al, 1986; Grammer & Thornhill, 1994; Jones & Hill, 1993), through studies of facial composites (Perrett et al, 1994; 1998; Rhodes et al in review) to the generation of attractive female face shapes using genetic algorithms (Johnston & Franklin, 1993) indicate that feminine features indicating estrogenised female faces increases their attractiveness cross-culturally. Sex typical female features (small lower face, a relatively flat mid-face, full lips and high eyebrows associated with a lack of brow ridge prominence) may indicate youth (as estrogen levels decrease in adult females with age), and reproductive health (Symons, 1979, 1984; Thornill & Gangestad, 1993, 1996, in press; Singh, 1993). A wealth of studies indicate that femininity (caused by an exaggeration of sex typical features) rather than averageness is attractive in female faces. Given the probable 'signalling' properties of estrogenised female faces, a male preference for such features is potentially adaptive.

There is some evidence for female preferences for exaggerated male facial characteristics. Scheib et al (in press) found a positive relationship between attractiveness and two markers of facial masculinity (cheek bone prominence and jaw size). Cunningham et al (1990) and Grammer & Thornhill (1994) used facial-metric measurements, and found a female preference for large jaws in males. 'Masculine' features, such as a large jaw and a prominent brow ridge are reliably associated with ratings of dominance in photographic, indenti-kit and composite stimuli (Berry & Wero, 1993; Berry & Brownlow, 1989; McArthur & Berry, 1987; McArthur & Apatow, 1983-1984; Perrett et al 1998). Facial dominance appears to correlate with status in some human hierarchies (Mueller & Mazur, 1997) and facial dominance in adolescent males is associated with earlier age at first copulation (Mazur et al, 1994). Nonetheless, the relationship between facial dominance and attractiveness is unclear – some studies find a positive relationship (Keating, 1985) while others find the opposite (Section 3.4, Berry & McArthur, 1985; McArthur & Apatow, 1983-1984; Perrett et al, 1998).

By asking subjects to rate black and white photographs, Berry & Zebrowitz McArthur (1985) found that baby-faces (non-masculine faces) were associated with elevated ratings of honesty, kindness and warmth (see chapter 3 for a fuller discussion of this, and similar studies). In addition, there was a significant correlation between babyishness and attractiveness ratings ($r=0.38$).

Berry (1991) had a large set of photographs rated for attractiveness, and independently, babyishness. She found, essentially, evidence for two kinds of attractiveness – sincerity (associated with attractive babyfaces), & power (associated with attractive 'mature' faces). As well as preferences for some typically 'masculine' features, Cunningham et al (1990) also found preferences for more feminine features (such as large eyes and an expressive mouth). Cunningham et al suggest that women have 'multiple motives' when choosing mates: optimally attractive faces may simultaneously appear dominant and yet elicit nurturing responses.

2.4.4.1 Manipulating sexual dimorphism in faces using computer caricaturing techniques

Although comparative approaches to human facial preferences suggest that exaggerated sex-typical traits will be attractive in both male and female faces, early studies (reviewed above) suggest that whilst this may hold true for female faces, preferences for male faces appear more complex. Work using computer graphic studies to manipulate facial sexual dimorphisms is introduced below and expanded upon and interpreted in chapters 3 and 4 below.

Computer graphic techniques can be used to construct 'average' male and female faces by digitally blending photographs of individuals of one sex (Figure 3; Chapter 7 for techniques). Sexual dimorphism in face shape can then be enhanced or diminished by looking at the geometrical differences between male and female face shapes and either exaggerating or decreasing them (Figure 4; Chapter 7 for techniques). For example, 'masculinising' a male face shape by increasing the differences between a male and female average increases the size of the jaw and reduces lip thickness (amongst all other dimorphic characteristics), as male jaws are larger than female jaws and the lips of men are thinner than those of women. These 'masculinised' or 'feminised' versions of male and female composite faces are useful stimuli for testing the influence of sexually dimorphic characteristics on judgements of facial attractiveness.

The amount of 'masculinity' or 'femininity' transform applied to a face is specified as a percentage throughout this thesis. The shape differences between male and female faces are described by a set of vectors between marked points on the features of the male and female averages (delineation points; sections 7.1.1 & 7.3.1). Transforms are expressed as a percentage of the distance traveled along these vectors: in a 25% 'feminised' male face shape, each delineation point is moved 25% of the way along the vector to the female average face. The colour information from the original male average is now warped into this new shape. To 'masculinise' male face shapes, the direction of the male-female vector is reversed before the points are moved along it.

Such manipulations of both Japanese and Caucasian face stimuli were presented to Japanese subjects in Japan and Caucasian subjects in Scotland (Perrett et al, 1998; Lee, 1998; Chapters 3 & 4 below). The technique was interactive: subjects could alter the appearance of a face on a computer monitor by using a computer mouse. Moving the mouse left or right would change the amount of masculinity or femininity in the face in real time, under the subjects' control (see section 7.5).

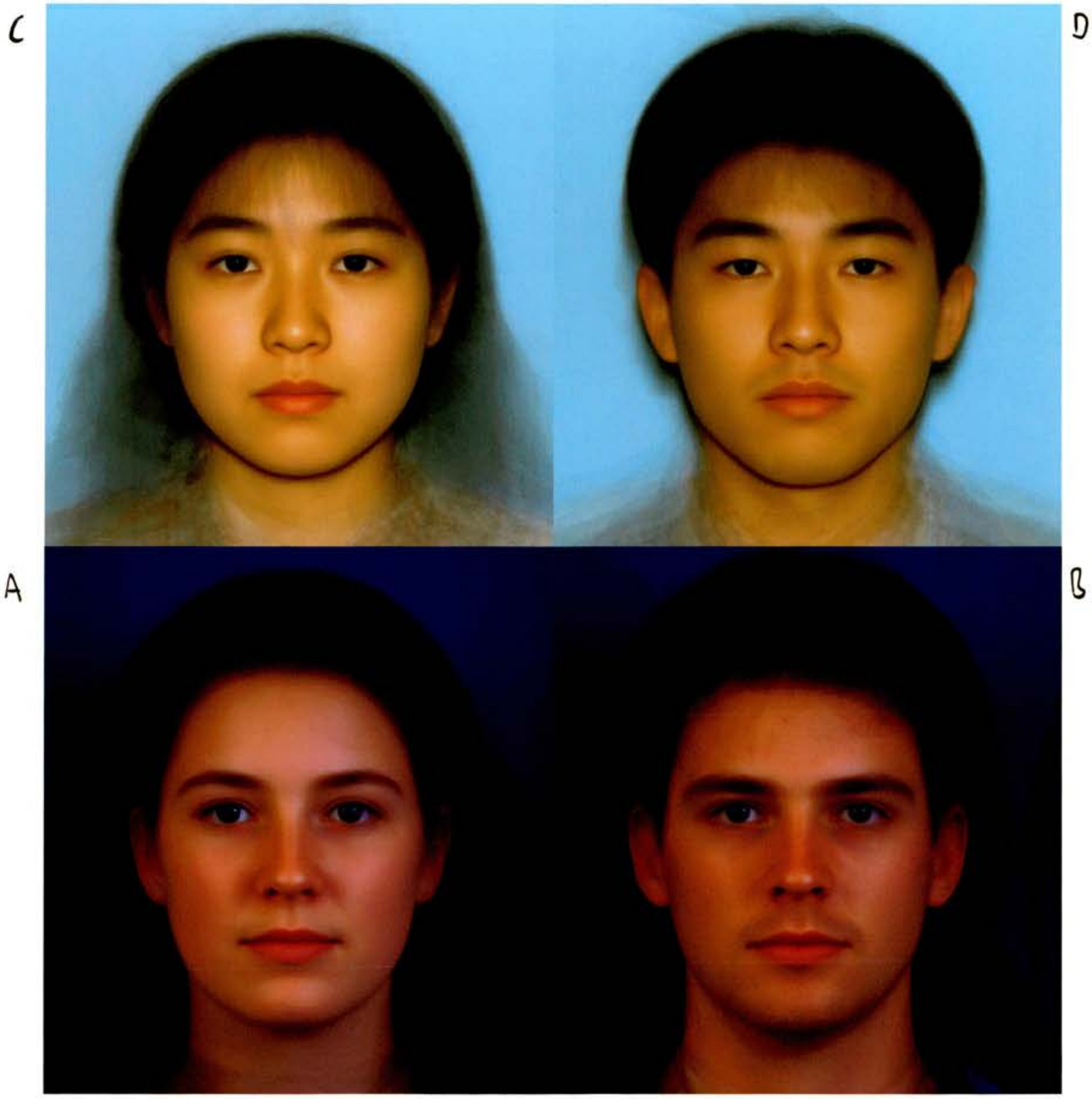


Figure 3: Composite 'average' facial images. 'Caucasian' female (a) and male (b) faces, 'Japanese' female (c) and male (d) faces. From Perrett et al 1998

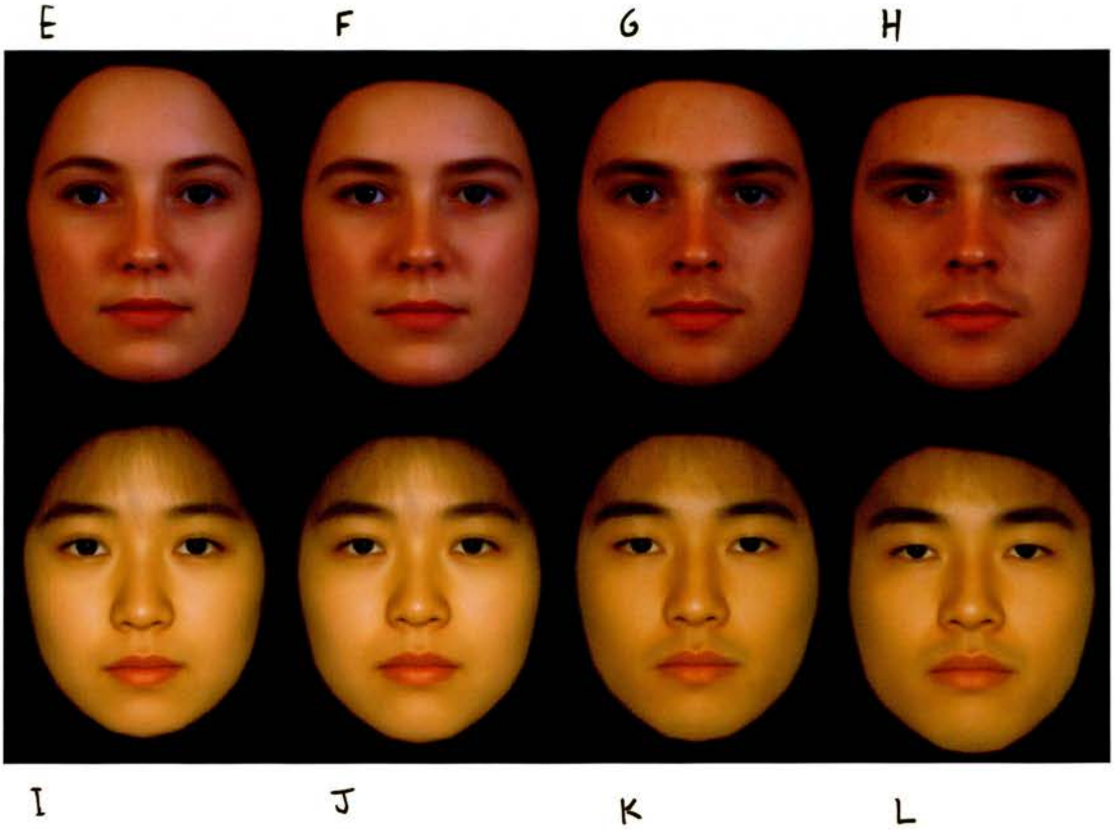


Figure 4: 'Feminised' and 'masculinised' facial images. Caucasian female, feminised 50% (e) and masculinised 50% in shape (f), Caucasian male, feminised (g) and masculinised (h), Japanese female, feminised (i) and masculinised (j) and Japanese male, feminised (k) and masculinised (l). From Perrett et al 1998

2.4.4.2 *Female masculinised and feminised stimuli*

The amount of masculinisation/feminisation transform applied by subjects to obtain the most attractive face shape was compared to a mean of 0% predicted by the null hypothesis (that altering sex-related characteristics would not effect attractiveness) and by the hypothesis that attractiveness is averageness (e.g. Langlois & Roggman, 1990). The face shape selected by Caucasian subjects as most attractive (from the shape range available) was significantly feminised for both the Caucasian and Japanese female face continua. Japanese subjects also selected significantly feminised versions of the female stimuli for both the Japanese and Caucasian female face continua (figure 5a). Neither the sex of the subject nor the race of the stimuli face influenced the results, but subjects preferred a greater degree of feminisation in their own population.

In summary, then, exaggerated sex typical features (small jaws, full lips etc.) appear to be attractive in female faces manipulated using this technique, consistent with other studies using various techniques reviewed above.

2.4.4.3 *Male masculinised and feminised stimuli*

For the male face stimuli, the shape selected by Caucasian subjects as most attractive (from the shape range available) was significantly feminised for both the Caucasian male face and the Japanese male face continua. Japanese subjects also selected significantly feminised versions of the male stimuli for both the Japanese and Caucasian male face continua. For the male stimuli, there was no effect of subject sex, subject population or type of stimulus face and no significant interactions between main effects (this data reported in Perrett et al, 1998 and Lee, 1998, figure 5b).

Previous studies show cross-population consistency in attractiveness judgements (Section 2.2 above). This study shows cross-cultural (between population) agreement in the preference for feminised to average face shapes which further refutes the averageness hypothesis. The study also indicates effects of experience on female attractiveness judgements since a greater degree of feminisation was preferred for faces from the subject's own population than for faces from a different population. Both generalisation and cultural specificity of attractiveness judgements may result from learning. Cues to female attractiveness relate to the way that female faces differ from males (see section 2.4.4 above). Sensitivity to the consistent sex differences in faces (and hence female attractiveness) could be learned through exposure to male and female exemplars. Most differences learned this way will generalise between populations as they reflect the common action of sex hormones during growth. Subjects, however, may become more sensitive to the sexual dimorphism of faces within the subject's own population because of increased exposure to population specific male-female variations.

This preference for 'feminised' male faces seems contrary to predictions from indicator mechanisms of sexual selection and to some other published studies of male facial attractiveness reviewed briefly above. However, other studies report a somewhat similar trend (e.g. Berry & Zebrowitz McArthur's (1985) reported preference for 'babyfaced' men). Rather than preferring typically masculine faces (with prominent brow ridges and large jaws), both male and female subjects appear to favour a small amount of femininity in men's faces, using stimuli prepared in this way. Chapters 3 and 4 below are an expanded account of data reported in Perrett et al (1998), and provide an interpretation of this seemingly anomalous result.

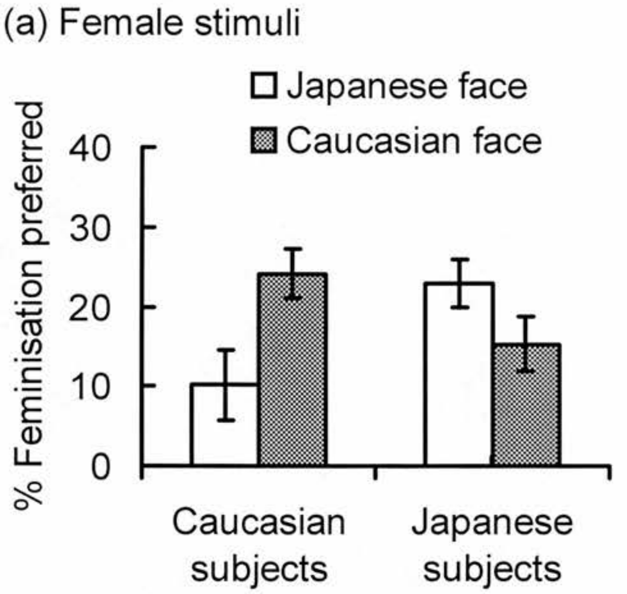


Figure 5a: The impact of feminisation of face shape on judgements of female attractiveness. Overall, subjects preferred a feminine face shape to an average shape both within and between populations. The degree of feminisation preferred was greater within than between populations for female faces. From Perrett et al 1998

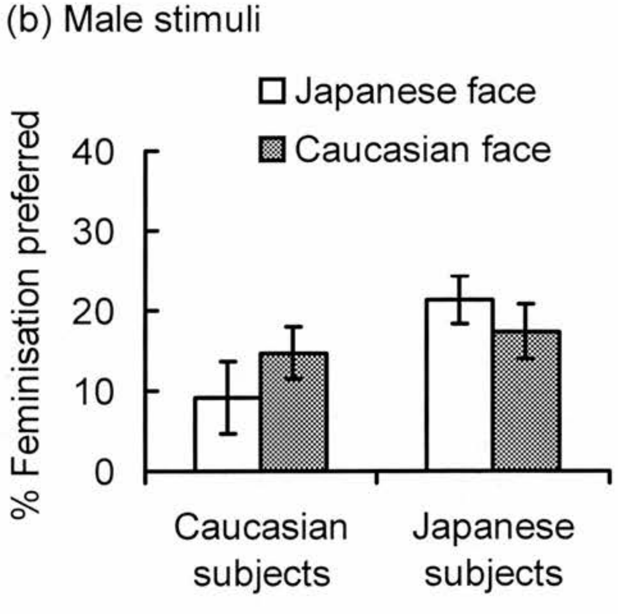


Figure 5b: The impact of feminisation of face shape on judgements of male attractiveness. From Perrett et al (1998)

3 Personality attributions and facial characteristics

3.1 Overview: a return to physiognomy?

In order to interpret the unexpected preferences for 'feminised' male faces reported above, the personality attributions these composite faces elicited were studied. The attribution of personality to faces, physiognomy, has had a long history across many cultures. Yet on the whole, scientists of the twentieth century have dismissed physiognomists as being, at best, misguided:

> 'Respectable science holds... that there is not any connection between the features of the face and the character of the person.... Any connection between these two wholly different kinds of personal qualities would bespeak of some mystical system of correspondences between the mind and body; scientists could only regard such a system as absurd' (Brandt, 1980; in Berry & Brownlow, 1989)

Despite such condemnation, Liggett (1974) demonstrated that the advice of science has clearly been ignored: 90% of university students see the face as a valid guide to a person's character. Despite the orthodox scientific view that physiognomy is a virtually worthless area for study that should be classified (and condemned) along with phrenology, recent literature contains a number of studies concerning personality judgements from physical characteristics. The review below has three main sections: a brief introduction to physiognomy since Aristotle, a review of studies indicating that observers can accurately judge individuals on at least some of the 'Big Five' personality factors from appearance alone, and finally a review of the literature on facial dominance. Finally an experiment is reported that shows high consensus in personality judgements attributed to masculinised and feminised composite face stimuli, providing an interpretation of the preference for feminised male faces reported in section 2.4.3.

3.2 A brief history of physiognomy

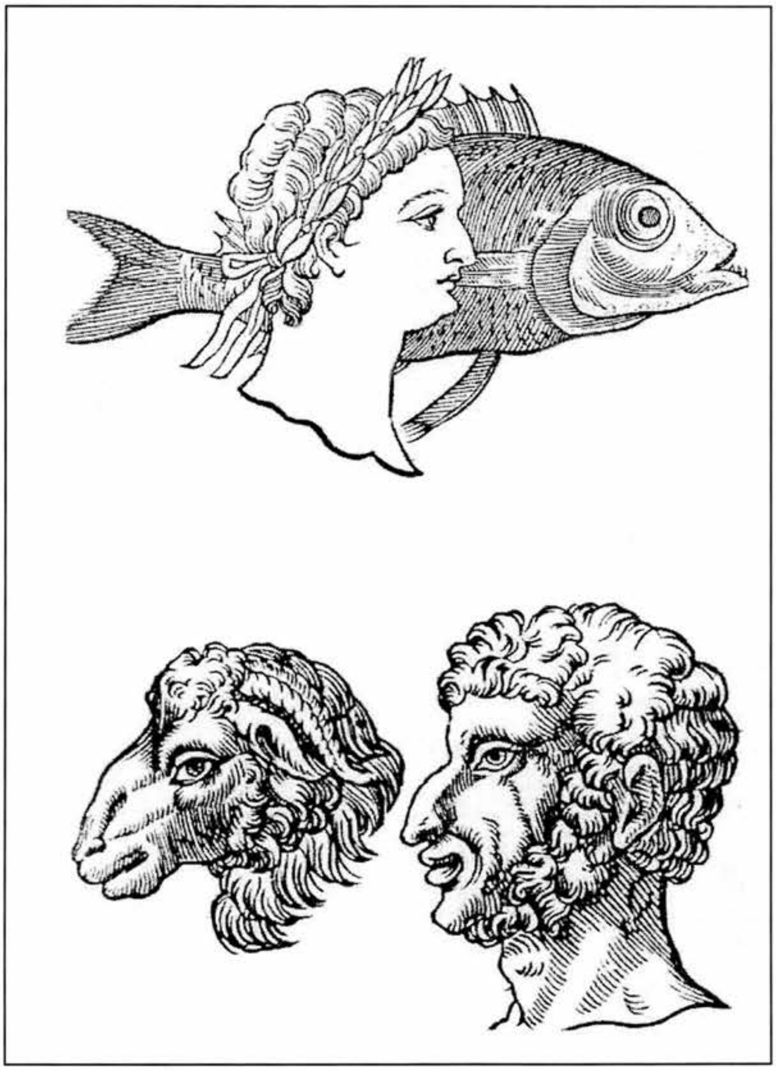


Figure 6: Comparisons of human and animal faces, by 16th century physiognomist Giovan Battista della Porta.

Physiognomy is a tradition over 2000 years old. The first written system of physiognomy – the *physiognomominca* – is attributed (probably incorrectly) to Aristotle (Berland, 1993). This work sets out three physiognomical methodologies, which can be summarised fairly easily (Berland, 1993). The first of these schools draws comparisons between humans and the characteristics of animals. For example, foxes are thought to be cunning, so people that have fox-like features are, likewise, cunning. This approach has no scientific justification – there is obviously no reason to suspect that inter-species comparisons can lead to valid personality judgements (Alley, 1988). Such comparisons have, however, led to some extremely imaginative illustration (Figure 6).

The second approach is based on racial or group differences and is again absolutely unscientific, revealing more about in-group/out-group prejudices than real measurable differences. Even the *physiognomica* does not favour this methodology. The third approach to physiognomy is to search for traces of emotional expression in faces. This methodology at least is based on the plausible assumption that the appearance muscular soft tissue may be altered by repeated use, or lack thereof. Muscle structure may offer some cue to the common affective state of an individual, if internal states are regularly shown in the face. There is even some evidence that spouses show convergence in their physical appearance over time, possibly as a result of shared emotional experiences over time (Zajonc et al, 1987; see also Chapter 6)

3.3 Studies of Zero Acquaintance and the 'Big Five' personality factors

Although often dismissed as charlatans or worse (e.g. Cleeton & Knight, 1924), the work of physiognomists may have been, in at least part, based on observable relationships between physical appearance and personality. Studies using 'zero acquaintance' paradigms (in which participants rate the personality of strangers) have found a surprising degree of correlation between self-ratings and stranger ratings on personality factors, often using five factor models of personality (e.g. Norman, 1963).

3.3.1 The five factor model of personality

Since Allport & Odbert (1936) constructed a list of 4,504 words in the English language that refer to characteristics of behaviour, personality theorists have argued as to the ideal taxonomy of trait dimensions underlying personality. Cattell (1966) refined Allport's extensive list and extracted 16 basic personality traits; Eysenck (1953), mirroring the Hippocratic system, proposed two major personality factors (extraversion and neuroticism - a third, psychoticism, was added later). Some consensus has been reached, however, and many trait theorists agree that five trait dimensions seems to represent a reasonable compromise that is at least a partially accurate measure of people's personalities (e.g. McCrae & Costa, 1987; Watson, 1989; Barret & Pietromonaco, 1997). These five dimensions have been derived from factor analyses of ratings of individuals using trait adjectives. Although different researchers disagree on the best names for these five factors, they are listed below with exemplar adjectives from various studies (McCrae & Costa, 1987; Norman, 1963; Botwin et al, 1997):

1. Extraversion (or Surgency): talkative-quiet, sociable-reclusive, dominant-submissive.
2. Agreeableness: irritable-good-natured, cold-warm, negativistic-co-operative.
3. Conscientiousness: Careless-careful, undependable-reliable, persevering-fickle.

4. Emotional Stability (or Neuroticism): Calm-anxious, even-tempered-temperamental, secure-insecure.

5. Openness-intellect (or Culture) ignorant-knowledgeable, boorish-refined, unadventurous-daring.

3.3.2 Studies of consensus at zero acquaintance.

The most relevant application of five-factor personality theory to this thesis is research that shows a surprising correlation between stranger-ratings and self-ratings of an individual's personality. Studies using 'zero acquaintance' paradigms in which subjects rate strangers show that humans are able to generate impressions of others' personality without social interaction. Not only do different judges generate similar responses to a given target individual, but also these impressions are more accurate than would be expected by chance (i.e. agree with the targets self-ratings). Whilst very few of the studies reviewed below specifically study the role of the face alone in these judgements, it seems likely that the face is an important cue to perceived personality.

In Passini and Norman's 1963 experiment, small groups of undergraduates were put together for a 15 minute period and asked to rate each other using scales that corresponded to the 'big five' personality factors, without verbal interaction. They found correlations between self and others' ratings of greater than 0.3 for 'Extraversion' and 'Conscientiousness', and 'Culture' factors. 'Agreeableness' (r=0.15) and 'Emotional stability' (r=0.02) failed to reach significance. Clearly, the 15 minutes that the group spent together provides opportunities for non-facial characteristics to influence personality judgements, such as clothing, posture and other non-verbal behaviour (eye-contact, smiling etc.).

Albright, Kenny, and Malloy (1988) replicated Passini & Norman's study, again using a paradigm that involved small groups of 4-6 individuals rating themselves and each other on personality dimensions. They found that self and others' ratings of the traits 'sociable' (extraversion) and 'responsible' (conscientiousness) were significantly correlated, but ratings of other personality traits were not. Albright et al found that the subjects' physical attractiveness positively correlated with ratings of the traits 'sociable', 'talkative' and 'good-natured', suggesting a halo effect. In addition, there was a strong relationship between raters' opinions of targets' neatness of dress, and ratings of conscientiousness, indicating that non-facial cues were employed to make judgements. Kenny, Horner, Kashy & Chu (1992) replicated Allbright et al's findings using videotapes of individuals rather than the small group paradigm, indicating that consensus at zero acquaintance is not due to interactions between the individuals rating each other.

Watson (1989) also replicated Passini and Norman's findings using a small group experiment, but allowed subjects to say one sentence introducing him/herself to the rest of the group. For individual judges, he found significant self-peer agreement for ratings of Extraversion and Conscientiousness, but no other factors. By averaging ratings across all judges that had rated a target individual, Watson also determined that increasing the number of judges increased the size of the self-peer correlations for Extraversion

($r=.43$, $p<0.05$), Agreeableness ($r=.31$, $p<0.05$) Conscientiousness ($r=.28$, $p<0.05$) and Culture ($r=0.20$, ns) but not Emotional Stability ratings ($r=-0.04$, ns).

Individual differences have been found in both accuracy of judging others, and of being judged oneself. Ambady, Hallahan & Rosenthal (1995) found that individuals who rated themselves highly on extraversion and agreeableness are more accurately perceived by others. They also report that women are more accurate judges of strangers' personality than men, consistent with other literature of non-verbal behaviour. Although individual differences in zero acquaintance judgements exist, Albright, Malloy, Dong et al (1997) demonstrate the cross-cultural nature of consensus of personality attributions. Using a small group paradigm with Chinese students in Beijing, they report a similar pattern of results as those found in Western studies. In a further study American participants judged personality traits of Chinese students from photographs, and vice-versa. They found that Chinese and American subjects showed much cross-cultural agreement in their judgements of personality at zero acquaintance.

Borkenau & Liebler (1992) attempted to identify the sources of validity for zero consensus judgements. Videotapes of target subjects reading a passage of text were made. Judges were presented with film with sound, silent film, a still photo or just the audiotape of the target reading. The still photograph condition (most relevant to the studies reported in this thesis) shows that even with extremely limited information and no transient facial expression, there were significant correlations between judges' and self-ratings of Extraversion ($r=0.33$) and Conscientiousness ($r=0.32$). Ratings of Agreeableness showing a non-significant but positive relationship ($r=0.19$).

Other studies essentially replicate the findings reported above (see Kenny, Albright, Malloy and Kashy, 1994 for a review) suggesting that from limited information (even still photographs) judges agree on the personality of other individuals. Furthermore, as these ratings appear to correlate reliably with self-rated extraversion and conscientiousness in all studies, and self-rated agreeableness and openness-intellect in a subset of the literature, they are to some extent valid. From the perspective of this thesis, it seems likely that stereotypical personality judgements made to faces may influence attractiveness judgements.

3.4 The effects of 'mature' and 'baby-face' characteristics on attributions of facial dominance.

Several researchers have devoted much effort to examining the characteristics and consequences of faces that are described as 'dominant' or 'baby-faced'. Dominance can be considered as a trait encompassed by the 'extraversion/surgency' factor in the five-factor model (although it appears to load more heavily on the 'agreeableness' factor in some studies, but it has received considerable attention in studies from outwith this theoretical orientation. Dominance is a characteristic that is closely linked to male reproductive opportunities and success in many species including humans. Dominant looking teenagers, for example, copulate earlier than less dominant looking adolescents (Mazur, Halpern & Urdy, 1994). The characteristics that lead to attributions of dominance may influence attractiveness judgements.

Most studies agree on the characteristics that make male faces appear dominant: 'Mature' features, such as a large jaw and a prominent brow ridge are reliably associated with ratings of dominance in photographic, indenti-kit and schematic stimuli (Berry & Wero, 1993; Berry & Brownlow, 1989; McArthur & Berry, 1987; McArthur & Apatow, 1983-1984). Similarly, the literature agrees that 'baby-faces' are characterised by smaller chins, high eyebrows and larger eyes. Such faces are generally rated as being warmer, more honest, and more sincere but also more naïve and less physically strong (Berry, 1991; Berry & Zebrowitz McArthur, 1986; Zebrowitz McArthur & Apatow, 1983-84). Baby-like and dominant faces reliably elicit personality attributions, but their effect on attractiveness judgements is equivocal, as discussed in section 2.3.3.

These perceptions of faces seem to generalise across cultures. Zebrowitz McArthur and Berry (1987) demonstrated that Korean subjects rated the personality of schematic Caucasian faces in the same way as American college students, indicating at least some cross-cultural agreement in these attributions. Stronger evidence for cross-cultural consensus in personality associated with faces comes from Keating, Mazur & Segall (1981). They found substantial agreement in dominance ratings of multiple race stimuli (European, African, Oriental-Polynesian & Indian (sub-continental) from participants in an extensive range of countries (Spain, Germany, Kenya, Zambia, Brazil, Columbia, China, Thailand and the United States).

The attributions made to a baby- or mature faced individual influence social interactions. Facial maturity has been shown to have more of an effect on male daily social experience than on female's routine interactions (Berry & Landry, 1997). By simulating a trial situation, Berry & Zebrowitz-McArthur (1988) demonstrated that manipulating the maturity of the 'defendant' influences both the chance of 'conviction' and the 'sentence' given. Baby-faced individuals were less likely to be found guilty of charges involving intentional criminal behaviour, although they were more likely to be found guilty of a 'negligent' crime. Baby-faced defendants were also given lighter sentences for 'negligent' crimes, reflecting the effects of attributed naivety and honesty. In a study with perhaps more ecological validity, Mueller & Mazur (1997) showed that facial dominance of the graduates from the West Point Military Academy in 1950 predicted their final rank at the end of their careers.

Consensus in judgements of the personality of mature or baby-like faces does not necessarily indicate that such attributions are valid, even though they influence potentially important social situations. Many of the personality attributions made to baby-like faces are also, unsurprisingly, made to infants themselves. Lorenz (1943) suggested that positive behavioural responses to infants are adaptive in that they increase offspring viability. Similar responses to adult faces with some childlike qualities may be inappropriate generalisations of an otherwise adaptive behaviour. This argument alone, however, fails to explain the evidence that baby-faced or mature individuals do to some extent self-report that they the personality characteristics attributed to them. Berry & Brownlow (1989) found that ratings of male babyishness were positively correlated with the face owner's self-reported approachability and warmth, but negatively related to self-reported aggression. For female faces, babyishness was associated with low self-reported levels of physical power and assertiveness. Bond, Berry & Omar (1994) demonstrated that individuals whose faces were rated as being 'less honest' were more likely to volunteer for experiments that involved them deceiving others than people who were judged to look more honest.

Such correspondence between self-reported and observed personality variables is often explained in terms of a behavioural confirmation or self-fulfilling prophecy effect (Berry & Finch Wero, 1993; Watson, 1989). It seems that certain facial configurations reliably elicit certain personality judgements. If someone is regularly perceived and treated as if they are submissive (or dominant), they may modify their behaviour so that they act in a more submissive (or dominant) manner. This, in turn, may lead to an internalisation of such behaviours, and an individual may come to see themself as possessing the characteristics that they are stereotypically attributed.

The behavioural confirmation theory of male strangers provides no explanation of the origin of stereotypes necessary for consistent ratings of dominance and other characteristics. Cultural factors are often invoked, and yet cross-cultural studies show considerable agreement in attributions of dominance (see above). This has led some researchers to posit a biological basis for the personality ratings given to faces. As the differences between 'dominant' and 'baby-faces' clearly parallel the differences between 'masculine' and 'feminine' composite stimuli, they probably reflect underlying androgen levels. Testosterone is also linked to male dominance behaviours (Mazur & Booth, 1998), tentatively providing a biological link between facial appearance, behaviour and valid (though stereotypical) personality judgements.

The behavioural confirmation theory and a putative biological theory of facial dominance may in fact interact and reinforce one another. Male testosterone levels are known to increase in 'winners' and decrease in 'losers' of competition (e.g.Mazur, Susman & Edelbrock, 1997). Success in dominance interactions may increase testosterone levels in a way that influences the growth of male facial characteristics, especially in adolescents. A reciprocal relationship between dominance behaviours influenced by stereotypical personality attributions and levels of testosterone is plausible (Gage, Manning, Scutt, Diver & Fraser, in preparation).

3.5 EXPERIMENT 1

3.5.1 INTRODUCTION

Given that people show a large amount of consensus when judging personality factors in real & schematic faces, an experiment was designed to see if there was consensus in judgements of the personality of masculinised and feminised composite stimuli. The literature reviewed above makes clear predictions: masculinised faces should be rated highest in dominance, but lowest in warmth and other traits relating to social approachability.

In self-report questionnaires, both males and females desire a kind and understanding partner more than they desire a physically attractive partner (Buss, 1989; Buss & Barnes, 1986). So, could stereotypical personality judgements influence attractiveness judgements? This is a particularly interesting question with respect to male facial attractiveness. Although the commonly accepted male physical ideal seems to fit a description of a more masculine than feminine face (square jaws etc.), subjects in earlier experiments picked a feminised composite as being the most attractive (section 2.4.4.3; Perrett et al 1998). Possibly, attributions of positive personality characteristics out-weigh classic male 'good looks' in attractiveness judgements.

Judgements of personality using bipolar trait adjective pairs are commonly measured with Likert scale ratings. Due to the similarity of masculinised and feminised versions of composite faces a ranking rather than a rating task was employed, to be more sensitive to differences in the attributions made to stimuli.

Faces were rated on thirteen traits in total, using descriptions which correspond well to factors identified as important for mate preference by independent research. Five of these trait adjectives can be found in 5 factor models of personality (Botwin et al, 1997; McCrae & Costa, 1987). They are listed here with the factors that they load on in McCrae & Costa's 1987 five-factor model:

Dominance (Extraversion and negatively on Agreeableness)

Warmth (Agreeableness)

Emotional (Neuroticism & Extraversion)

Intelligence (Conscientiousness & Openness/Intellect)

Co-operativeness (Agreeableness)

Three further personality traits were selected. Berry & Wero's (1993) study indicated that raters could accurately identify a stranger's willingness to indulge in deceptive behaviour: the adjective '*Honesty*' was used to reflect this ability. The trait '*Assertiveness*' was suggested by an anonymous referee to reflect properties of 'Dominance' but with, perhaps, fewer negative connotations. The concept of a '*Good Parent*' reflects a composite of several personality traits rather than a discrete characteristic of an individual. It was included to investigate whether participants showed consensus in attributing family centred behaviours to faces that reflected high or low androgens.

The remaining five characteristics reflected other (mainly physical) characteristics that may possibly be attributed to the stimuli. *'Masculinity'* of the stimuli was ranked to see if our computer graphic manipulations of male-female differences in facial structure accurately altered subjective impressions of masculinity. As masculinised stimuli have more mature facial characteristics than feminised stimuli, they were ranked on perceived *'Age'*. High androgen levels are associated with honest advertisement of immunocompetence in males (section 1.4) but contraindicate reproductive health in women, so perceived differences in *'health'* between the stimuli were investigated. Stimuli were ranked for height (*'Tall'*) to see if physical size is positively associated with masculinity in males or females. Finally, as dominance is associated with hierarchical status in men, stimuli were ranked for *'wealth'*.

3.5.2 METHOD

To interpret the preferences for 'feminised' facial images reported above, 50% masculinised, 50% feminised and average images (Figure 2) were rated for perceived characteristics by a new set of subjects. Twenty subjects from St Andrews (age range 18-50, 10 female) were presented with 4 sets of 3 images representing the end points of each continuum and the average. Subjects were asked to rank stimuli from one set on 7 characteristics (masculinity, dominance, warmth, emotionality, honesty, intelligence and age). The order of testing of characteristics and image sets was randomised. An additional 20 subjects (age range 19-61, 10 female) ranked the stimuli on 6 further characteristics (co-operativeness, assertiveness and 'good parent', wealth, health and height).

3.5.3 RESULTS

For Caucasian and Japanese male faces, increasing the masculinity of face shape across the 3 set members increased ranking of perceived dominance, masculinity and age but decreased ranking of perceived warmth, emotionality, honesty, co-operativeness and quality as a parent (Friedman's $\chi^2 \geq 15.6$, df=2, $p < 0.0005$, for each rated dimension). Increasing masculinity affected the two female face sets in the same way for all characteristics ($\chi^2 \geq 8.1$, df=2, $p < 0.017$, each dimension) except 'good parent' with the Caucasian female faces where the rank order was average, feminised and masculinised ($\chi^2 = 6.7$, df=2, p=0.035). Increasing masculinity did not consistently decrease apparent intelligence (Caucasian male and female faces, $p > 0.5$, Japanese female face, p=0.07, Japanese male face, p=0.02) or increase attributions of assertiveness (Japanese and Caucasian female faces, $p > 0.5$, Japanese male face, p=0.058, Caucasian male face, p=0.157). No consistent rankings were found for the health dimension (Japanese faces, ns; Caucasian male faces, feminised preferred, $\chi^2 = 17.2$, df=2, $p < 0.003$; Caucasian females, average preferred, $\chi^2 = 14.8$, df=2, $p < 0.001$). 'Wealth' and 'height' attributions were at chance levels for Japanese faces and Caucasian males, although feminised Caucasian female faces were ranked as being wealthiest and tallest ($\chi^2 = 7.5$, df=2, $p < 0.024$; $\chi^2 = 14.7$, df=2, $p < 0.001$ respectively). These results are shown in full in tables 1-4.

Caucasian male faces	Mean Rankings			χ^2 (N=20) 2 d.f.	P
	Fem.	Ave.	Masc.		
Dominant	1.10	1.90	3.00	36.4	.000
Warm	2.60	2.25	1.15	22.9	.000
Emotional	2.65	2.10	1.25	19.9	.000
Intelligent	2.05	2.15	1.80	1.3	Ns
Masculine	1.40	1.70	2.90	25.2	.000
Honest	2.65	2.25	1.10	25.9	.000
Old	1.25	1.80	2.95	30.1	.000
Assertive	1.85	1.80	2.35	3.7	Ns
Co-operative	2.60	2.10	1.30	17.2	.000
Healthy	2.45	2.15	1.40	11.7	.003
Parent	2.35	2.25	1.40	10.9	.004
Tall	2.15	2.05	1.80	1.3	Ns
Wealthy	1.95	2.20	1.85	1.3	Ns

Table 1: Rank order of personality attributions made to Caucasian male faces

Caucasian female faces	Mean Rankings			χ^2 (N=20)	P
	Fem.	Ave.	Masc.	2 d.f.	
Dominant	1.65	1.80	2.55	9.3	.010
Warmth	2.55	1.90	1.55	10.3	.006
Emotional	2.70	1.95	1.35	18.3	.000
Intelligent	2.15	2.00	1.85	0.9	Ns
Masculine	1.15	1.95	2.90	30.7	.000
Honest	2.55	2.10	1.35	14.7	.001
Old	1.60	1.85	2.55	9.7	.008
Assertive	1.95	1.90	2.15	0.7	Ns
Co-operative	2.60	2.10	1.30	17.2	.000
Healthy	2.30	2.40	1.30	14.8	.001
Parent	1.90	2.45	1.65	6.7	.035
Tall	2.55	2.10	1.35	14.7	.001
Wealthy	2.50	1.75	1.75	7.5	.024

Table 2: Rank order of personality attributions made to Caucasian female faces

Japanese male faces	Mean Rankings			χ^2 (N=20) 2 d.f.	P
	Fem.	Ave.	Masc.		
Dominant	1.20	2.10	2.70	22.8	.000
Warmth	2.50	2.20	1.30	15.6	.000
Emotional	2.65	2.15	1.20	21.7	.000
Intelligent	2.20	2.30	1.50	7.60	.022
Masculine	1.10	2.05	2.85	30.7	.000
Honest	3.00	1.90	1.10	36.4	.000
Old	1.35	1.80	2.85	23.7	.000
Assertive	1.65	1.95	2.40	5.7	0.058
Co-operative	2.60	2.15	1.25	18.9	.000
Healthy	2.22	2.05	1.73	2.6	Ns
Parent	2.65	2.05	1.30	18.3	.000
Tall	1.92	2.00	2.08	.228	Ns
Wealthy	1.75	1.95	2.30	3.1	Ns

Table 3: Rank order of personality attributions made to Japanese male faces

Japanese female faces	Mean Rankings			χ^2 (N=20) 2 d.f.	P
	Fem.	Ave.	Masc.		
Dominant	1.50	1.75	2.75	17.5	.000
Warmth	2.60	2.10	1.30	17.2	.000
Emotional	2.85	2.10	1.05	32.7	.000
Intelligent	2.10	2.30	1.60	5.2	Ns
Masculine	1.35	1.90	2.75	19.9	.000
Honest	2.85	1.85	1.30	24.7	.000
Old	1.40	1.95	2.65	15.7	.000
Assertive	2.10	1.80	2.10	1.2	Ns
Co-operative	2.50	2.25	1.25	17.5	.000
Healthy	2.35	1.95	1.70	4.3	Ns
Parent	2.45	2.00	1.55	8.1	.017
Tall	2.20	2.00	1.80	1.6	Ns
Wealthy	2.10	2.05	1.85	0.7	Ns

Table 4: Rank order of personality attributions made to Japanese female faces

3.5.4 DISCUSSION

The results of this experiment demonstrate several points. Firstly, exaggerating or reducing sexual dimorphism in composite faces alters their perceived masculinity or femininity in the predicted direction. Manipulations of this sort clearly influence perceptions that convey the 'psychological meaning' of masculinity and femininity (Meyer & Quong, 1999).

Secondly, there is considerable consensus in the ranking of all stimuli across most traits tested. Although masculinised and feminised versions of stimuli are fairly similar, they must contain stereotypical cues to personality. 'Socially valued' traits such as honesty, warmth, co-operation and skill as a parent are associated with feminised versions of male faces, whilst traits such as 'dominance' are, as predicted, associated with masculinised face shapes.

Finally, some traits that theory predicts to be associated with masculinised male faces, such as height, health and wealth (through the action of testosterone on physique and dominance behaviour) are not reliably linked to masculinity or femininity in male faces. Indeed in Caucasian male faces, health is linked to feminised faces in direct opposition to the immunocompetence principle, which is perhaps interesting with respect to Kalick et al's (1988) finding that attractiveness judgements of faces do not predict long term health of their owners in contemporary society.

A female preference for slightly feminised male faces can perhaps be at least partially explained by the results of this experiment. Although biological predictions such as handicap theory indicate that masculinised faces should be preferred, such faces elicit negative personality attributions. Cross-culturally, personality factors are reported to be the most important factor in mate choice by both sexes (Buss & Barnes, 1986; Buss, 1989). It seems inconceivable that personality attributions have no effect on attractiveness judgements. As reviewed above, personality attributions, though stereotypic, may predict behaviour. For example, ratings of perceived dishonesty from facial appearance correlate with the face owner's willingness to participate in deceptive behaviour (Bond, Berry & Omar, 1994). Feminisation of male face shape may increase attractiveness because it 'softens' particular features that are perceived to be associated with negative personality traits.

3.5.5 Testosterone, behaviour & face shapes

The putative link between testosterone, face shape and behaviour outlined briefly above deserve further discussion. A possible positive feedback loop between social interaction, testosterone and face shape was outlined above: males who succeed in dominance encounters may, as a result, have increased levels of testosterone which may be reflected in the growing face. Increasing testosterone level in adult males is associated with more troubled sexual relationships. A sample of 4462 former members of the US armed forces showed that men with testosterone levels one standard deviation above the mean (estimated from one measurement) are 50% more likely to never marry than men with testosterone one standard deviation below the mean (Booth & Dabbs, 1993). The same study demonstrated that married men with high testosterone were also more likely to suffer troubled relationships, and demonstrated increased incidence of domestic violence and extra-pair sex.

The implications of these findings for theories of sexual selection in humans are clear: whatever indirect 'good genes' benefits high testosterone may offer in terms of offspring viability, there is an apparent cost in terms of likely paternal investment. As an aside, it is interesting to reiterate that women make more accurate personality judgements in zero-acquaintance studies than men (Ambady et al, 1995). As males lower initial investment in offspring makes them the sex most likely to desert (Symons, 1979), selection pressures to make accurate personality judgements are stronger in females than males.

In the holistic transforms employed to masculinise and feminise the stimuli used in this experiment, all male-female facial differences visible in frontal facial photographs are manipulated. This methodology has generated a robust and reliable preference for slightly feminised male stimuli using subjects (university students) matched in age with those stimuli, both in the UK and Japan. This preference is unexpected and at odds with some of the other literature on male facial attractiveness and studies of sexual selection in other species, as reviewed above. It is also contrary to reported male sexual behaviour. Dominant looking teenage males, for example, copulate earlier than their less dominant looking peers (Mazur et al, 1984). Facial dominance influences rank in human military hierarchies (Mueller & Mazur, 1997) and 'cultural success' as reflected by status in social hierarchies is positively correlated with copulation frequency in contemporary Canada (Perusse, 1993). Cues to cultural success (such as high status clothing) undoubtedly positively influence attractiveness judgements of males (e.g. Townsend & Wasserman, 1998; Townsend & Roberts, 1993). Testosterone appears to be related to dominance behaviours and so, by extension, position in heirarchies, but simultaneously testosterone is associated with less likelihood of extended paternal investment (Mazur & Booth, 1998).

One aspect of female mate choice may be a cost-benefit analysis of such factors: a high status/testosterone partner may have access to more resources, and offer heritable benefits to offspring. Such a partner, however, may be more likely to desert or even injure the female who chooses him (Mazur & Booth, 1998). A female preference for a lower status partner who is more likely to provide long term paternal investment may, overall, pay a greater reproductive dividend. Given that stereotypical personality judgements have some accuracy and women make more accurate attributions than men, such attributions may form an important component in women's mate choice decisions. In our stimuli, personality attributions apparently outweigh biological fitness cues in attractiveness judgements of male faces. In real faces, perceived personality may not be as closely linked to masculine characteristics: combinations of masculine and feminine features, indicating both status and kindness may be optimally attractive. There is some support for such 'multiple motives' influencing judgements of facial attractiveness (Cunningham et al, 1990; see also Scheib & Gangestad, in press).

Sexual dimorphism in any species reflects compromises amongst diverse selection pressures. In humans, the greater upper body musculature and more rugged skeletal anatomy of males relative to females may reflect advantages in male-male competition and hunting. Since male attractiveness is an important determinant of relationships and sexual partnerships, the reduction in attractiveness of male face shape with masculinisation represents a further selection pressure. This would act against 'run away' Fisherian sexual selection for extreme male characteristics and is consistent with the relative lack of sexual dimorphism in humans.

The preferences found here indicate a selection pressure on the evolution of face shape that acts against dramatic differences between males and females and, since more feminine face shapes are perceived as younger, the preferences would encourage a youthful, neotonous appearance in the species generally. There is some palaeo-anthropological evidence consistent with such a hypothesis, reviewed below.

3.5.6 Perceived age and a preference for femininity of male faces

The preference for slightly feminised male face shapes may reflect the effects of masculinity on perceived age. While status and height are valued in males (Buss, 1989; Jackson, 1992), youth benefits attractiveness judgements for both female and male faces (Jones, 1995; Deutsch et al, 1986). For both males and females, enhancing sexual dimorphism in face shape develops cues to characteristics which, from a biological perspective, appear beneficial (i.e. youth and fertility in females and dominance in males).

Meyer (1999) suggests that the subjects in Perrett et al may be attracted to youth cues rather than femininity in faces. Cues to femininity and youth are, however, inextricably related due to sex differences in the amount of facial growth at puberty. Whilst pubescent male faces grow a great deal, female faces change less and therefore remain relatively neotenous. As we have shown, manipulations based on geometric sex differences in *face shape alone* affect apparent age. Certainly, face stimuli derived from males with measured high and low levels of testosterone would reflect androgenic effects on target tissues more accurately than our stimuli. Nevertheless, high testosterone composite faces will carry more 'mature' features and may still appear older than low testosterone composites.

Evolutionary theories do not predict that females should favour youthful looking males, instead they predict that cues to high male status, such as dominance, will be considered attractive (Buss, 1989). Our masculinised male faces were rated as more dominant than feminised or average facial stimuli. Facial dominance is a predictor of male status in at least some human dominance hierarchies (Mueller & Mazur, 1997) and yet subjects in our experiments preferred sub-dominant male faces. Whether this is due to a preference for youthfulness, or a preference for other characteristics attributed to femininity (such as apparent warmth, kindness or a willingness to invest in progeny), is unclear. What is apparent is that explanations of male attractiveness based on female sensory biases (Enquist & Ghirlanda, 1998) or honest Zahavian handicaps fail to predict the preference for slightly feminised male facial characteristics indicating that females select mates on characteristics additional to or other than dominance.

3.5.7 Decreasing robusticity in the fossil record.

Although a preference for more gracile male faces seems opposed to much recent theorising in human sexual selection, the fossil record provides plentiful evidence that *homo* is becoming less robust. The facial and skeletal characteristics of ancestral hominids became less robust with passing time (figure 7; Aiello & Dean, 1990; Ruff, Trinkaus & Holliday, 1997). There is some evidence that the trend has been continuing within recent history. Since the beginning of the last period of glaciation (100, 000 years ago) until 10,000 years before the present, tooth size decreased by 1% every 2,000 years. In the last 10,000 years however, a relaxation of selection pressures due to the development of food preparation techniques has led to a doubling of the rate of tooth size reduction (Brace, Rosenberg & Hunt, 1987). The mechanism behind dental reduction is still debated with natural selection and genetic drift mechanisms proposed (e.g. Brace et al, 1987; Calcagno & Gibson, 1988, 1989; Sciulli & Mahaney, 1991). The evidence of a preference for feminised (gracile) face shapes may provide a possible role of sexual selection in this still debated process.

Figure 7: Decreasing facial robusticity in the fossil record? A series of skulls from the genus *homo,* arranged in chronological order. Facial robusticity appears to have been decreasing across hominid evolution. It is important to note, however, that not all these specimens are ancestral to modern *Homo sapiens.*

Such changes are not limited to facial and dental characteristics. From analysis of post-cranial fossils (mainly limb bones), Ruff, Trinkaus, Walker & Larsen (1993) conclude that:

> 'There has been a consistent, exponentially increasing decline in diaphyseal robusticity within Homo that has continued from the early Pleistocene through living humans. Early modern H. sapiens are closer in shaft robusticity to archaic H. Sapiens than they are to Recent humans.'

Ruff et al attribute these processes to technological and cultural advances that made heavier musculature and accompanying skeletal robusticity unnecessary. The accompanying decrease in natural selection pressures is hypothesised to have led to the dental and post-cranial changes, rather than sexual selection for gracile forms. Nonetheless, sexual selection theories that above average male muscle mass signals superior genetic quality (though the action of testosterone) and is therefore selected (either through male-male competition or female choice) receive little support from palaeoanthropological evidence. Sexual selection has been suggested as a possible cause of the trend for increasing body size in fossil mammalian lineages (Andersson, 1994, pg. 248). Yet body mass has neither increased nor remained constant in *h. sapiens:* it has actually decreased, despite advances in nutrition. As body mass is a correlate of testosterone (Thornhill & Gangestad, 1996) an evolved preference for feminised male faces is consistent with the fossil record, in a way that a preference for masculinised faces would not have been.

4 Stimuli Artifacts and Judgements of Attractiveness

4.1 Factors in stimuli construction

The stimuli used in experiment 1 above (Lee, 1998; Perrett et al, 1998) and illustrated in figure 2 raise concerns that must be addressed before firm conclusions concerning a preference for feminised male faces can be drawn. These concerns centre on two aspects of the stimuli: (a) with respect to sexual dimorphism in face size, and (b) placing a mask around the face, which is of dubious ecological validity and may introduce asymmetries and irregularities in the facial outline. A further criticism suggested by Meyer & Quong (1999), is that our stimuli may not represent biological masculinity and femininity at all. Below, the nature of these concerns is discussed in some detail, and an experiment to address them is presented.

4.1.1 Masculinisation, feminisation and sexual dimorphism in face size.

An adult male face is bigger than an adult female face, yet in the stimuli illustrated in Figure 4 (above) the Caucasian female face appears longer than the Caucasian male face. This is due to the standardisation procedure used to match the inter-pupil distance of the male and female composite images before the shape differences were amplified to masculinise or feminise the images (see chapter 7, section 7.3.1). In effect, this normalisation process involves either 'shrinking' the male prototype or 'stretching' the female prototype so that the eye centres are equivalent. For male faces the eyes lie closer to the face outline (compared to female faces), therefore standardising the eye positions increases the size of the female face outline relative to the male face. The facial proportions, however, remain valid; female faces are more childlike than males and have eyes closer together relative to the outline of the face, a larger forehead and a smaller chin relative to the rest of the face, all of which can be seen in the stimuli. When the differences between stimuli are exaggerated, however, the weakness of this normalisation procedure becomes apparent: masculinising a male face actually makes it smaller. Thus, a preference for a feminine face shape is confounded by size in the Lee (1998) data set, as although the configuration of the features is more feminine, the overall face size is larger: a male trait. This phenomenon is illustrated in figure 33 in section 7.3.1.

4.1.2 Masking the hair, ears and neck of masculinised and feminised stimuli

The masking technique used in Lee (1998; see chapter 7.2 for details) may introduce apparent inconsistencies in the masking of the stimuli. Two aspects of masking are problematic: (1) a question of ecological validity, and (2) the introduction of asymmetries in facial outlines (Figure 2 above)

The outline of the face stimuli is defined by straight lines connecting 26 points marking the average hair and chin outline (Chapter 7.1.1 & 7.2). The masks used in the study are made by removing the background around this outline and replacing it with a black mask. Blurring across the mask edge is necessary to avoid making the image look like a many sided polygon – effectively smoothing the edges. Unfortunately, this process does not translate well from the Silicon Graphics computer workstations to printer and so the problems of masking are exacerbated by problems of reproduction. The figures illustrated here have a clear, angled outline and a change of surface sheen that was not present during stimuli presentation.

Masking, nonetheless, can make the faces less realistic and can introduce apparent asymmetries into the facial outline, but there are theoretical and practical arguments in favour of masking facial stimuli. Technically, graphic techniques for accurately modeling and averaging hair are still unavailable. The averaging process we employ leaves hair looking somewhat blurred and unnatural (see Figure 1 and stimuli in Perrett et al 1994). Subjects are instructed to arrange hair so that it does not cover the face, nonetheless the hair may occlude parts of the forehead, ears and neck. Masking has the benefit of removing cues from hairstyle. Male and female hairstyles can differ quite dramatically. Therefore, manipulating male-female differences exaggerates differences in hairstyle (e.g. feminising increases hair length). Masked stimuli may be superior for experimentation on facial shape because they allow interpretation without preferences attributed simply to spurious fashions in hairstyle (e.g. preference for long hair). Moreover, the neck and ears are not consistently sampled due to differing neck-lines of clothes and hairstyles. Exaggeration of male-female differences introduces visual differences in the shape of the neck, which may not reflect actual sexual dimorphism.

The irregularities in the hairlines visible in figure 2 are a result of directional asymmetries in the sample of faces studied. Whilst the features of average composite faces tend to symmetry, hairstyles are less co-operative. It appears that in both the Caucasian and Japanese male samples, individuals were more likely to part their hair on the right-hand side of their heads, leading to occlusion of the left-hand forehead. This asymmetry is exaggerated when faces are masculinised, but reduced when faces feminised: feminised faces, then, have smaller apparent asymmetry in Lee (1998), and hence may be rated as more attractive.

4.2 EXPERIMENT 2

4.3 INTRODUCTION

The following experiment explores the effects of masking (mask versus no mask) and normalising techniques (retaining or reversing naturally occurring size dimorphism) on the amount of masculinity or femininity preferred in male and female faces. Such a study may go some way to answering the criticisms of the stimuli used in Lee (1998) outlined in sections 3.1.1 & 3.1.2 above.

4.3.1 METHOD

STIMULI

A new set of Caucasian male and female faces (n=21, 41 respectively, mean age 20 years) were photographed and male and female set averages constructed (face set 3: IPV_TW, see section 7.1.4.3 for photographic details). From these composites, 50% masculinised and 50% feminised male and female composites were generated. Two sets of masculised/feminised male and female averages were produced: one set using the technique employed by Lee (1998), which reverses naturally occurring size dimorphism (section 3.1.1 above; section 7.3.1), and another set constructed with a new technique that maintained naturally occurring size dimorphism (section 7.3.2 for techniques).

Masked and unmasked versions of each of these four masculinised/feminised face pairs were produced (section 7.2). Interactive continua were then produced from each of the face pairs. Interactive techniques of stimuli presentation allow the subject to manipulate the characteristics of the stimuli in real time. In this experiment subjects could alter the amount of masculinisation or feminisation in the face by moving a computer mouse left and right. The left-right movement causes a smooth morph of the face to any point on the continua between 50% masculinised and 50% feminised versions of the stimuli. Hardware texture mapping in Silicon Graphics Maximum Impact workstations allows intermediate images to be calculated on the fly. This technique should be very sensitive to subjects' preferences. For clarity, the eight continua used in this experiment are detailed in table 5 below.

SEX OF STIMULI	STIMULI MASKED?	WITH NATURAL SIZE DIMORPHISM?
MALE	YES	YES
MALE	YES	NO
MALE	NO	YES
MALE	NO	NO
FEMALE	YES	YES
FEMALE	YES	NO
FEMALE	NO	YES
FEMALE	NO	NO

Table 5: The eight continua employed in experiment 1. All continua ran from 50% feminisation to 50% masculinisation.

SUBJECTS

61 students at St. Andrews University (34 female, 27 male, mean age 21.3) participated. Sessions were self paced and lasted approximately 10 minutes.

PROCEDURE

Participants altered each continuum interactively in real time until the face on the screen was as attractive as possible from the shape range available. Stimuli were presented in a 600*600 pixel frame. Each continuum was presented twice to allow left/right counterbalancing, making sixteen continua in total. Order of continua presentation was randomised. Responses were recorded as a percentage of the transform applied to the face on screen when it was judged most attractive.

4.3.2 RESULTS

Responses to male and female stimuli were analysed seperately, below:

Male stimuli:

A 2*2 repeated measures ANOVA, with presence/absence of mask and presence/absence of size dimorphisms as factors and subject sex as a between subjects factor was performed.

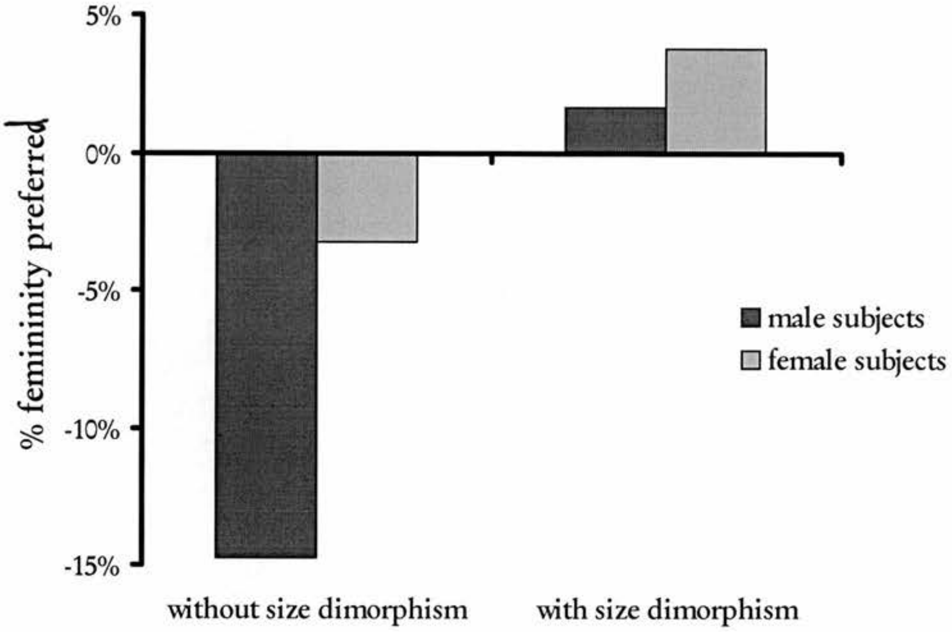
Results showed a main effect of sexual dimorphism: when natural size dimorphisms were maintained in stimuli significantly more feminisation was preferred in male stimuli ($F_{(1,59)}=16.2$, $p=.0001$; figure 8)

Masking had no significant main effect ($F_{(1,59)}=3.3$, $p=0.075$), although a trend indicated that masking may reduce the amount of feminisation. There was no significant effect of subject sex on responses ($F_{(1,59)}=2.2$, $p=0.14$), despite the apparent differences illustrated in figure 8.

All second and third order interactions were non-significant with F values lower than one, with the exception of the dimorphism*masking interaction ($F_{(1,59)}=3.42$, $p=0.069$): a trend for masking to have more effect on responses to dimorphic stimuli.

The absolute levels of masculinity/femininity preferred in each version for the composite presented varied greatly. Subjects significantly feminised faces in only one of the conditions (unmasked stimuli with sexual dimorphism. Mean response = 7% feminised: one-sample t-test against a hypothesised mean of 0%, $t_{60}=2.51$, $p=0.015$). In both of the stimuli with no sexual size dimorphism however, a significantly masculinised facial configuration was preferred (with mask: mean = 8% masculinised, $t_{60}=2.36$, $p=0.022$. Without mask: mean = 8% masculinised, $t_{60}=2.69$, $p=0.009$).

Male stimuli: effects of size dimorphism



Male stimuli: effects of masking

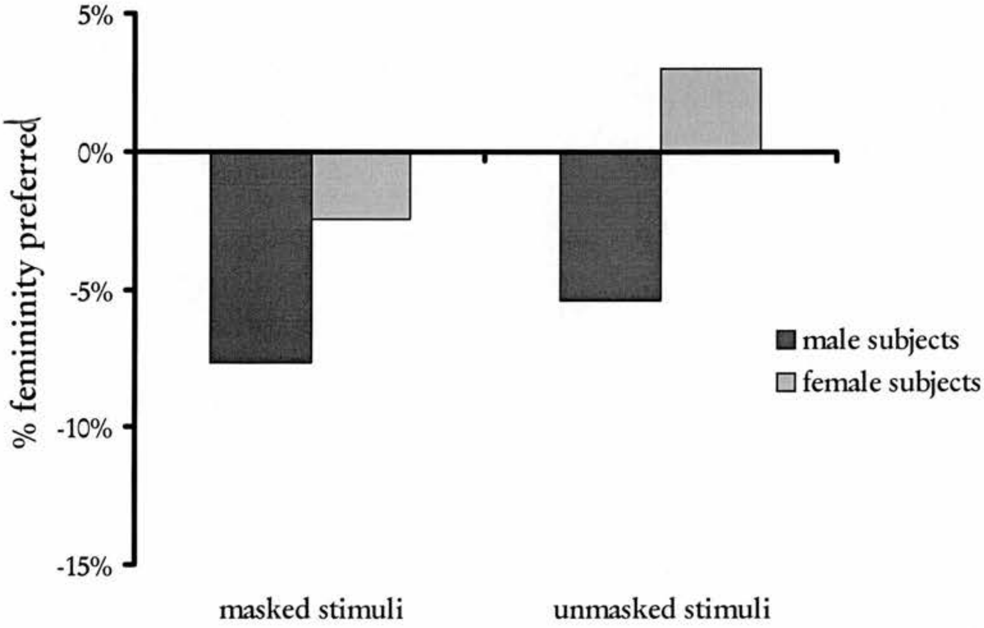


Figure 8: Effects of size dimorphism (top) and masking (bottom) on femininity preferred in male stimuli. Positive values equal a preference for feminised faces, negative values indicate a preference for masculinised faces.

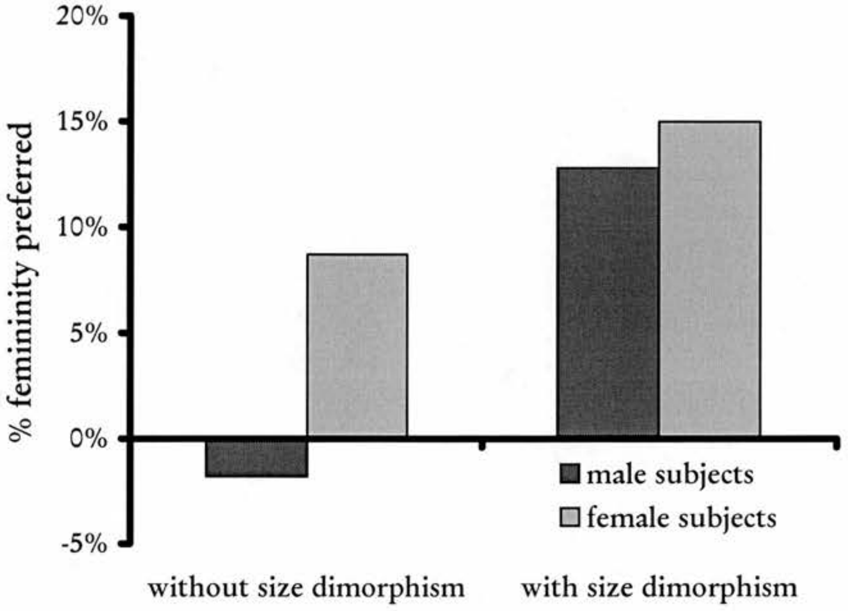
Female Stimuli

For female stimuli, another 2*2 repeated measures ANOVA, with presence/absence of mask and presence/absence of size dimorphisms as factors and subject sex as a between subjects factor was performed.

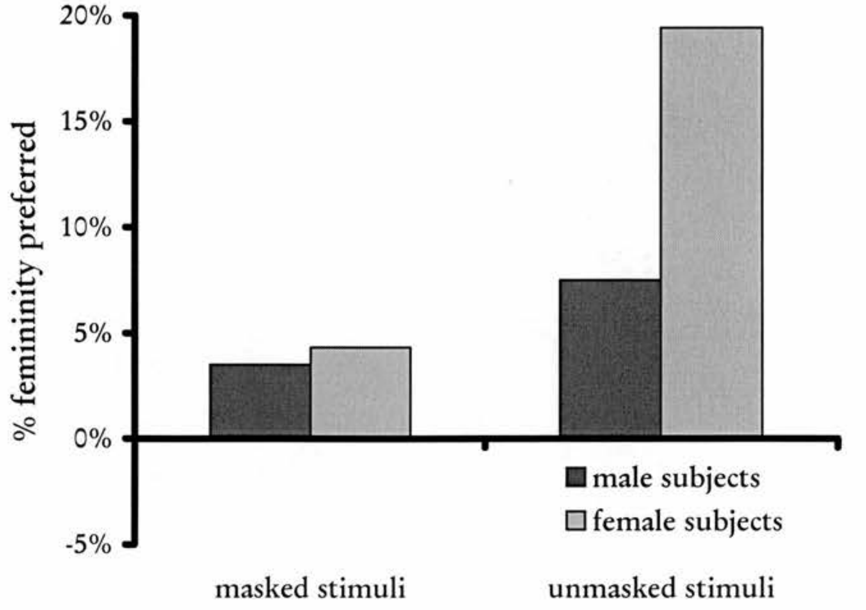
As with the male stimuli, a main effect of sexual dimorphism indicated that stimuli with accurate size information were feminised more than those without ($F_{(1,59)}=13.136$, $p=.0001$; figure 9 top panel) .

A main effect of masking ($F_{(1,59)}=14.2$, $p=0.0001$) indicated that the subjects preferred a higher level of femininity in unmasked rather than masked stimuli.

Again there was no main effect of subject sex $F_{(1,59)}=2.2$ $p>.1$, but a mask by subject sex interaction ($F_{(1,59)}=4.8$, $p=0.032$), indicated that masking has more influence on females' responses than males' responses (Figure 9). A Tukey HSD test indicated that male and female responses to unmasked female stimuli without sexual dimorphism were significantly different ($p=0.017$), whereas male and female responses to all other stimuli were statistically indistinguishable using post hoc measures. All other interactions had Fs < 1.



Female stimuli: effects of size dimorphism



Female stimuli: effects of masking

Figure 9: Effects of size dimorphism (top) and masking (bottom) on femininity preferred in female stimuli.

4.3.3 DISCUSSION

These results reported above answer questions concerning masking and dimorphism: both seem to decrease the amount of feminisation preferred in these composites, and probably in the Lee (1998) study. So, the experiment reported in Perrett et al (1998) probably *underestimates* the preference for feminised face shape. The effects of masking and dimorphism seem to be complex, however, and may depend to some extent on the sex of both the stimuli and subject. However, it is disconcerting that absolute levels of femininity are not equal to those in Lee (1998).

In the condition directly comparable to the Lee (1998) study (masked stimuli without sexual dimorphism), no preference for femininity was found for male or female subjects. In fact, male subjects preferred a significantly masculinised male face in this condition.

Female subjects only preferred a feminised male face in one of the four continua presented (unmasked with natural levels of dimorphism) whereas males appeared to judge a masculinised face shape as attractive in the two continua without natural sexual dimorphism. The apparent lack of a preference for femininity in this experiment is important, and raises questions about the nature of 'average' composite stimuli. This point will be returned to in detail in section 5.5 (experiment 6), which reports a study in which stimuli generated from five face sets were presented to subjects, allowing a better analysis of differences between stimuli.

4.3.4 The Bio/logic of facial geometry?

Meyer & Quong (1999) raised one criticism of the stimuli generated by caricaturing male/female differences. They note that the masculinised and feminised stimuli used in Lee (1998), Perrett et al (1998) and in this thesis represent both 'geometrical' masculinity and femininity (as they are constructed from shape differences between male and female stimuli) and 'psychological' masculinity/femininity (as deduced from ratings studies - chapter 4 below). They question, however, whether the masculinised stimuli really represent faces with higher levels of androgens, arguing that such faces should be generated from same sex individuals with measured high or low androgens.

Visible sexual dimorphism in the human face increases rapidly at puberty when sex steroids influence tissue growth (see section 2.3 above). The working assumption behind these techniques of stimuli generation is that the geometrical differences between male and female faces parallel the differences between individuals with high and low androgens. Masculinity in mammals is testosterone dependent whereas femininity represents the more neutral developmental state (Owens & Short, 1995). Individuals with an XY chromosome pattern and complete androgen insensitivity syndrome have a female appearance (although they do have male dentition Haaq & Donohoe, 1998). Whilst stimuli constructed from individuals with known levels of androgens and oestrogens would be superior to the technique employed here, it seems inconceivable that geometrically 'feminising' a face produces facial attributes associated with *higher* than average levels of androgens (Perrett & Penton-Voak, 1999).

5 Female Sexual Preferences, Behaviour and the Menstrual Cycle

5.1 Concealed ovulation and the human female menstrual cycle

In contrast to most other mammals and many other primates, human female reproductive cycles do not have a clearly defined estrous period of sexual activity around ovulation. Instead ovulation is hidden both from males and to (most) females themselves. Women can and do have sex at any point in their menstrual cycle, leading some theorists to view human sexual behaviour as 'emancipated' from hormonal influences (Ford & Beach, 1951). More recently, many researchers have shown great interest in the possibility that constant receptivity is an adaptive behaviour, allowing varied female reproductive strategies. In this chapter, theories of concealed ovulation are reviewed, and then evidence that some mate choice relevant female preferences may change across the menstrual cycle, possibly to facilitate mixed reproductive strategies. Finally, three experiments that investigate the effects of menstrual cycle hormonal shifts on female preferences for male faces are presented.

5.1.1 Evolutionary theories of concealed ovulation in humans

Several competing evolutionary explanations have been proposed to explain concealed (or cryptic/hidden) ovulation. Early 'social bonding' hypotheses proposed that with division of labour (hunting males & gathering females) continual sexual receptivity would strengthen the human pair-bond, and hence co-operation within the group necessary for successful hunting/gathering lifestyles (e.g. Lancaster, 1975; Morris, 1968). Recent ethnographical and archaeological studies suggest that this may be unlikely (see section 1.4.3).

Alexander & Noonan (1979) suggested that ovulation was concealed as an adaptation to 'force desirable males into consort relationships', and thereby increase paternal investment. As ancestral males had little or no idea when a female was ovulating they were coerced into an extended period of contact with a female to improve chances of conception. In turn, this extended companionship would increase the paternity certainty of the male whilst reducing his chances of seeking other sexual partners. Alexander & Noonan (1979) use parental investment theory (Trivers, 1972) and propose that in proto-humans such a strategy could tip male reproductive strategies toward increased paternal investment and monogamy, and away from the search for multiple partners. However, as numerous commentators point out, if male parental care increases reproductive success, it will be selected for: males do not need to be 'tricked' into pair bonds if they are associated with greater reproductive success (Burley, 1979; Blaffer Hrdy, 1979).

Two related theories are proposed by Benshoof & Thornhill (1979) and Blaffer Hrdy, (1979). Benshoof & Thornhill consider that ovulatory crypsis coupled with the impossibility of round the clock mate guarding allows females to cuckold their primary mates with partners who may offer some advantage (either genetic or otherwise). They propose that hidden ovulation evolved after, rather than before, ancestral hominids adopted social monogamy as a mating system. Ambiguous paternity (a given male can never be completely sure that a woman's offspring is his own) may have prevented infanticide in ancestral multi-male primate groups (Blaffer Hrdy, 1979, 1981). Hrdy's hypothesis for the loss of estrous is unique as it assumes hidden ovulation evolving in a non-monogamous group. Schröder (1993) expands and adapts the infanticide/cuckoldry hypothesis. Noting that copulation in humans in all cultures normally takes place in private, she suggests that concealed ovulation with a tendency for clandestine copulation may have allowed females to cuckold dominant males in ancestral uni-male groups.

Burley (1979) uses ethnographic evidence in support of a radically different, yet fascinating, idea. Her theory proposes that with the evolution of hominid consciousness came a gradual realisation that pregnancy and childbirth were both painful and dangerous to females: increasing brain size in *homo* infants coupled with a bipedal gait make birth relatively more dangerous than for other primates. Hence, hominid females may have avoided copulation around ovulation to avoid getting pregnant: Burley's review of attitudes to childbirth in less developed contemporary societies supports her theory: men want more children than their partners, who often want to limit family size. Burley suggests that women who were less aware of ovulation may have been less successful at avoiding conception and hence bore more children as a result; an outcome favoured by natural selection. This hypothesis fits with data concerning the unreliability of female self-reports of menstrual cycle activity: women seem surprisingly poor at reporting dates of previous menstrual periods (e.g. Bean et al 1979; Baker & Bellis, 1995). Recent studies, however, indicate more accurate self-report (Baker et al, 1998; Small, 1996).

5.1.2 Concealed ovulation in non-human primates.

Most theories of human concealed ovulation (which date, for the most part, from the late 1970s) start from the twin assumptions that humans are almost unique in having concealed ovulation and that this evolved form an ancestral state in which hominids advertised ovulation through estrous. In fact neither of these assumptions is supported conclusively by comparative data.

Concealed ovulation is found in a substantial minority of primates (47% of 68 anthropoid taxa reported by Sillen-Tullberg and Møller , 1993; 24 out of 78 primate species studied by Blaffer Hrdy & Whitten, 1987). Very obvious estrous swellings are found in only 22 of around 200 total primate species. Of these 22 species most are macaques and baboons (Small, 1993). Small (1993, 1996) considers that comparative analysis with our closest relatives (Bonobos, who have near permanent sexual swellings, and chimpanzees, who have estrous) precludes certainty of the ancestral state: humans may not have lost obvious ovulation. Certainly, the sexual swellings found in chimpanzees are derived from a common ancestor without such obvious advertisements (Burt, 1992; Dixson, 1983).

Despite recent findings, Sillen-Tullberg and Møller (1993) consider that the common ancestor of *pan* and *homo* had slight signs of estrous. Their analysis indicates that cryptic ovulation evolved 0-1 times in primates with a monogamous mating system, and 8-11 times in non-monogamous contexts, depending on the phylogeny employed. Nonetherless, monogamy has become the mating system 4-6 times subsequent to the loss of advertised estrous and just 1-3 times in other primate species.

The comparative data available now casts doubt on some of the original theories concerning the origin of hidden ovulation. Female primate receptivity across reproductive cycles is not exclusively linked to monogamous pair bonding, threatening Morris' pair-bond theory (Blaffer Hrdy, 1981; Small, 1993). The number of primate species who have cryptic ovulation indicates that the sophistication of Burley's conception avoiding hominids is not a necessary precursor to cryptic ovulation.

Benshoof & Thornhill's theory that estrous was lost to allow cuckoldry in a socially monogamous primate ancestor is also questionable; cryptic ovulation evolved either once or (more likely) never in already monogamous primate species (Sillen-Tullberg & Møller , 1993). Hrdy's conception that concealed ovulation would help to prevent infanticide in multimale groups is, however, consistent with comparative and phylogenetic analyses, although not with fossil evidence indicating large sexual dimorphism in ancestral hominids which may reflect a unimale social structure.

Overall there appears to be some confusion as to the evolutionary origins of concealed ovulation. Hypotheses of the origin cryptic ovulation are often very similar to those concerning sexual swellings: Hrdy (1981) considers both cryptic and advertised ovulation may allow females to mate with multiple males to introduce paternity uncertainty (see also Dixson, 1983). Primate species do seem more likely to become socially monogamous in the context of cryptic ovulation (Sillen-Tullberg & Møller , 1993). Whatever events led to the lack of estrous (be it an adaptation or the result of some stochastic process), once such a loss had occurred processes such as those suggested by Alexander & Noonan become possible. Sexual strategies involving cuckoldry may also become easier to execute by females once ovulation has been lost (e.g. Benshoof & Thornhill, 1979). Some evidence seems to support such a theory in humans, reviewed below (section 5.2.2).

5.2 Changes in sexual behaviour across the menstrual cycle

Although women are considered constantly 'sexually receptive', many researchers have proposed that female sexual desire, behaviour and preferences may vary across the menstrual cycle. The sources of this variation may be social (e.g. sex during menstruation and post-partum is taboo in many cultures; Hrdy, 1981) or biological (resulting from the influence of fluctuating levels of sex hormones). Before reviewing the evidence for changing female sexual behaviour across the menstrual cycle (section 5.2.2 below) I will briefly review the biology of the human female cycle.

5.2.1 Hormonal changes across the menstrual cycle.

Women's menstrual cycles last between 21 and 35 days, with a mean duration of around 28 days. Many commentators have pointed out the great irregularity of cycle lengths both within and between women, which is especially interesting in comparison to the regularity of most mammals (e.g. Burley, 1979). Nevertheless, most standard models of the menstrual cycle are based on a duration of 28 days. In such models, ovulation occurs on approximately day 14 (Regan, 1996).

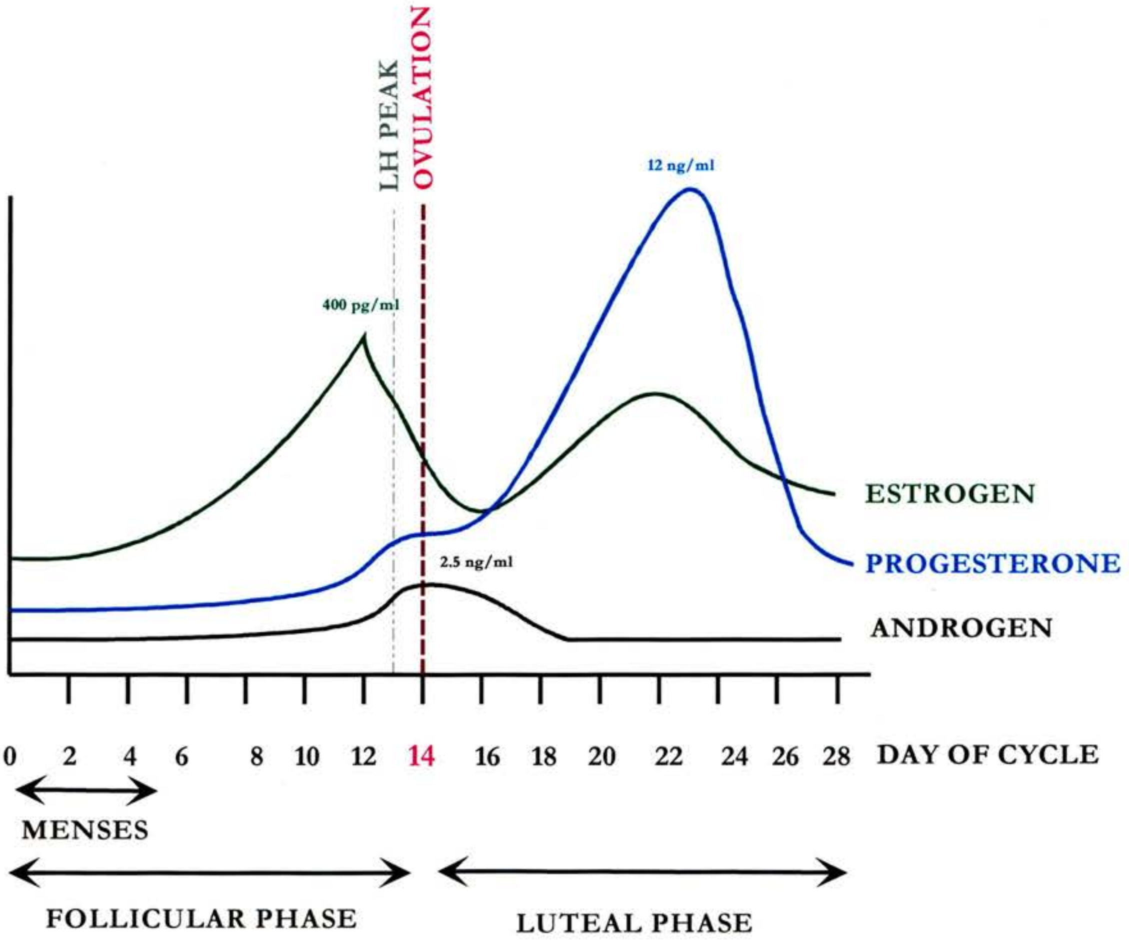


Figure 10: Hormonal changes across the menstrual cycle. Not to scale. Peak hormone levels are shown for estrogen, progesterone and androgen (adapted from Regan, 1996)

During the fourteen days prior to ovulation, an ovarian follicle, containing an ovum, develops under the influence of follicule stimulating hormone (FSH) and lutenizing hormone (LH). This follicular phase can be sub-divided into three shorter phases. The first five days (early follicular phase) is the period of menstruation, in which levels of estrogen, progesterone and androgen are all low. In the mid-follicular phase (days 6-10) estrogen levels rise steadily, androgens increase somewhat and progesterone levels remain low. Days 11-14 (the late follicular phase) are characterised by a rise in androgen, progesterone and a rapid increse of estrogen levels, which peak approximately 2 days before ovulation. This estrogen peak causes a surge of LH, which, in turn leads to the follicle to rupture and the ovum to be released on day 14 in standard models.

The fourteen days following ovulation, the luteal phase, are characterised by the formation of the corpus luteum (which produces progesterone) from the ruptured follicle. In this phase, progesterone levels increase to a peak on around day 23, and estrogen levels increase also (see figure 10). A drop in both estrogen and progesterone levels occurs before the onset of menstruation and the beginning of the next cycle. A review of current understanding of the neuroendocrine regulation of the menstrual cycle by the ovary, pituary and hypothalamus can be found in Chabbert Buffert et al, 1998.

As human sperm survive in the female reproductive tract for around 5 days (Baker & Bellis, 1995), copulations in the five days preceding ovulation (mid- to late follicular phase) are most likely to result in conception. Once ovulation has taken place, copulation is unlikely to lead to pregnancy. Despite the model given above, ovulation is actually fairly unpredictable. There is even the possibility that women, like some other mammals, may ovulate in response to copulation during the follicular phase (Jochle, 1973; Baker & Bellis, 1995).

5.2.2 Changes in sexual behaviour and preferences across the menstrual cycle

Accumulating evidence suggests that some aspects of female sexual preference and behaviour change across the menstrual cycle with the probability of conception. Despite so-called sexual receptivity across all of the menstrual cycle, several researchers hypothesise that desire and frequency of sexual activity may vary across the menstrual cycle.

Peaks in sexual desire and activity have been reported around ovulation (e.g. Urdy & Morris, 1968; Adams, Burt & Gold, 1978; Stansilaw & Rice, 1988); in the mid-follicular phase (Urdy & Morris, 1977; Bancroft, Saunders, Davidson, & Warner, 1983), and also in the late luteal/pre-menstrual period (e.g. Stewart, 1989). Further studies report multiple peaks of sexual desire and/or activity, often at ovulation and in the late luteal phase (e.g. Silber, 1994). Overall, then, the picture of fluctuating sexual activity and desire is confused: dissociating various factors in research is problematic (e.g is sex initiated by the female or male partner? Is fluctuating sexual interest the result of hormonal changes, or societal pressures against sex during certain phases of the menstrual cycle?). Regan (1996) presents a comprehensive review of the literature and concludes that there is substantial variation between women who may have no, one or two peaks of sexual desire. Women that do have peaks in sexual desire tend to have them either in the mid-follicular, late follicular (ovulation) or late luteal phases of the menstrual cycle. Follicular phase or mid-cycle peaks in sexual interest would parallel many other primate species, as such peaks would be associated with likely conception following sex (Hrdy, 1981). Luteal phase peaks in sexual activity, however, cannot directly help conception.

Despite confusion as to exactly when peaks in sexual interest occur, two studies report that women may be more inclined to seek sex outside the pair bond when conception following sex is most likely. Baker & Bellis (1990, 1995) report data indicating that the rate of female extra-pair copulations (EPCs) is around 2.5 times higher during the follicular phase than in the luteal phase, possibly to promote sperm competition when conception is likely. A decrease in EPC frequency during non-fertile phases of the menstrual cycle may reduce the risk of EPCs being discovered, and hence the risk of desertion by the cuckolded male. Baker and Bellis' study has generated much controversy as the data was collected from a questionnaire in 'Company' magazine, introducing the possibility of some self-selection bias amongst participants. A further study (Worthman, 1978; reported in Hrdy, 1981) supports the findings in a non-industrial setting, using hormonal measurements to determine menstrual status. Worthman reports that there is a significant peak in mid-cycle sexual activity with both husbands and lovers amongst women of the !Kung San.

Some evidence also suggests that female preferences for male characteristics may change across the menstrual cycle. Fitness benefits to offspring can only be realised if conception follows copulation, so females may be more attentive to phenotypic markers indicating fitness during the follicular phase of the menstrual cycle when conception is most likely (days 6-14; Gangestad & Thornhill, 1998; Regan, 1996; Baker & Bellis, 1995).

Female preferences for odour seem to have a cyclic component, with females being more receptive to odours at ovulation, coincident with increases in estrogen and lutenizing hormone (Doty, Snyder, Huggins & Lowry, 1981). Odour cues appear to play an important role in human mate choice, and women especially seem to be able to discriminate between males on characteristics that are fitness relevant on the basis of smell (e.g.Wedekind, Seebeck, Bettens & Paepke,1995; Wedekind & Furi, 1997).

Grammer, (1993) studied female perception of androstenone, a chemical found mainly in male rather than female sweat. In contrast to the closely related androstenol, which has a pleasant, sandal-wood like smell that is attractive to females, androstenone has an unpleasant, urine-like odour. The exact mechanism of androstenone & androstenol production is unclear: they may both be by-products of enzymes action on bacteria, or androstenol may oxidise to androstenone (see Grammer, 1993, for a brief review). Grammer, 1993, showed that the generally unpleasant odour of androstenone becomes more acceptable to females mid-cycle in women who do not use hormonal contraception: a preference for a dimorphic trait at ovulation.

Two other studies of women's olfactory preferences across the cycle show that women may also make discriminations of male 'quality' in the follicular phase of the menstrual cycle that they do not make at other times (Gangestad & Thornhill, 1998; Thornhill & Gangestad, in press). Both of these studies employed similar methodologies in which females rated the odour of t-shirts that had been worn by males for two nights. Previously, the level of fluctuating asymmetry of these men had been estimated from measurements of digit length, elbow width, wrist width, ankle width, foot breadth and ear length and width. Both experiments demonstrated that women (not using hormonal contraception) favoured the odour of symmetrical males, but only in the follicular phase of the menstrual cycle. Further more, the more recent study showed that women prefer the scent of men whose facial photographs are rated as attractive at times of high fertility. Men, on the other hand, showed no preference for the odour of symmetrical females (Thornhill & Gangestad, in press).

This cyclic change can be interpreted from a 'good genes' perspective. Symmetry is a marker of (possibly) heritable developmental stability (sections 1.3.4 above) and is probably a correlate of secondary sexual trait size and hence high levels of androgens (Scheib et al, in press; Thornhill & Gangestad, 1996). A preference for symmetric men when conception is most likely may be adaptive, increasing offspring viability. In combination with other evidence, such a mechanism may be especially important in extra pair sexual activity. Females prefer symmetrical men as extra-pair copulation partners, and evidence reviewed above indicates that EPC activity peaks in the follicular phase (Thornhill & Gangestad, 1994; Gangestad & Thornhill, 1997; Baker and Bellis, 1995). As symmetry appears to be inversely related to the investment that men make in relationships, such choices probably do not reflect selection for paternal effort (Thornhill & Gangestad, in press; Gangestad & Thornhill, 1997a, 1997b).

Finally, a similar cyclic change has been shown to operate in the visual modality, Frost (1994) developed photographs of male faces so as to manipulate the skin darkness. Skin colour is sexually dimorphic within all races, with males having darker skin than females (Van Den Bergh & Frost, 1986). Frost's study showed that females prefer darker photographs of male faces around ovulation. Given the sexual dichromatism of skin colour, a preference for darker skin at ovulation could be considered a preference for exaggeration of a male trait at ovulation.

5.3 EXPERIMENT 3

5.3.1 INTRODUCTION

Female sexual behaviour and preferences change across the menstrual cycle, favouring male 'quality' when conception is most likely. The findings for a preference for feminised faces (Perrett et al, 1998) was interpreted with respect to the positive personality attributions that such faces received: such faces are perceived as being more likely to provide direct phenotypic benefits.

The findings of previous work (section 5.2.2 above), however, indicate that sexually dimorphic traits may become more attractive at high fertility phases of the menstrual cycle. Frost's evidence for cyclic visual preferences for male faces in females suggests that a study of cyclic female preferences for male facial shape may be fruitful. It is hypothesised that females may be attracted to more 'masculine' faces with exaggerated secondary sexual facial characteristics when conception is most likely than at other times of the menstrual cycle.

5.3.2 METHOD

STIMULI

The stimuli employed in this experiment were based on the photographs of Caucasian males (Set 3, IPV_TW) described in section 7.1.4.3. Two feminised male face shapes were generated (representing 30% and 50% of the vector differences between male and female face shapes). Two 'masculinised' face shapes were constructed by exaggerating the vector differences between points on the male and female averages by 30% and 50%. Identical colour information from the average face was then warped into the feminised and masculinised faces. Masculinisation and feminisation of the stimuli reflected naturally occurring size dimorphism in male and female faces: masculinised faces were larger than the feminised faces (section 7.3.2). Manipulating male-female differences exaggerates differences in hairstyle (e.g. feminising increases hair length). This sexual dimorphism reflects fashion more than anatomy, so stimuli were masked around the outline of the face to prevent such cues influencing attractiveness judgements (section 7.2). All five stimuli are illustrated in Figure 11.

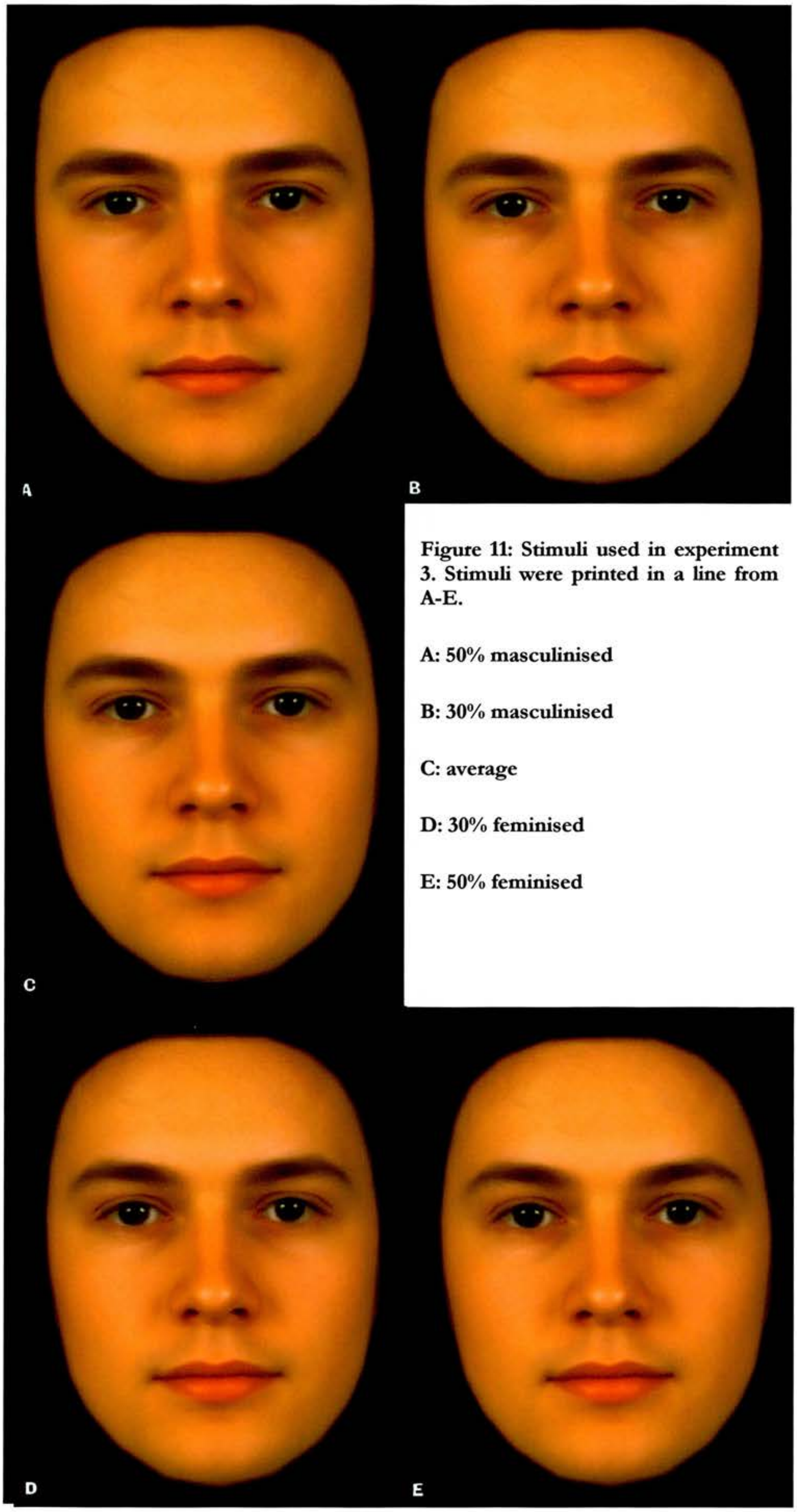


Figure 11: Stimuli used in experiment 3. Stimuli were printed in a line from A-E.

A: 50% masculinised

B: 30% masculinised

C: average

D: 30% feminised

E: 50% feminised

SUBJECTS AND PROCEDURE.

The five stimuli (50 & 30% masculinised, the average, and 50 & 30% feminised) were printed (each face: 65 by 85mm) in full colour in the Tomorrow's World Magazine, a national UK monthly published by the BBC. A short accompanying questionnaire asked respondents which of the five stimuli they considered 'most attractive' and requested details of age, oral contraception use, pregnancy and number of days since the onset of previous menses. Subjects returned the questionnaire to a freepost (pre-paid postage) address. Subjects that failed to complete details of their menstrual cycle, or who reported more than 28 days since the beginning of their last cycle were excluded from the analysis, leaving 178 completed questionnaires.

As cyclic odour preferences have not been found in women using oral contraception, women reporting pill use were excluded from the initial analysis (n=39). Following Gangestad & Thornhill (1998), a standard 28 day model of the female menstrual cycle was used to assign the 139 remaining female respondents (mean age of 30.7 years with a range of 14-50) to one of two groups based on their chance of conception. This was estimated from the number of days since the onset of the participants last menses: 'high conception risk' (the follicular phase, days 6-14) or 'low conception risk' (days 0-5 and 15-28, menses and the luteal phase: Baker & Bellis, 1995; Regan, 1995).

5.3.3 RESULTS

Figure 12 shows the percentage of subjects that selected each face in the low and high conception risk groups. Responses in the high (n=55) and low (n=84) conception risk groups were not normally distributed (One-sample Kolmogorov-Smirnov Z=2.10 and 2.07 respectively, $p < 0.001$ in both cases), so non-parametric tests were used in all analyses.

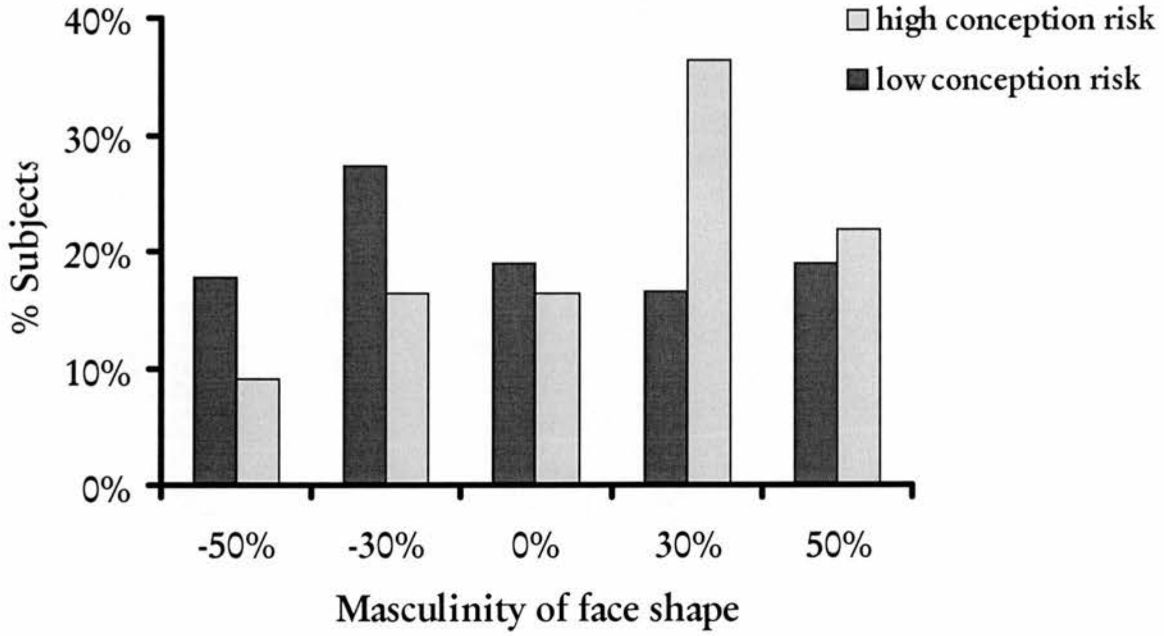


Figure 12: Percentage of subjects in high and low conception risk groups selecting each of the five stimuli as most attractive.

A Chi-squared test demonstrated that the high conception risk group were not selecting faces at chance levels (χ^2=11.45, 4 d.f., p=0.022). Inspection of Figure 12 shows a clear preference for 30% masculinised faces in this group. Despite the apparent preference for 30% feminised faces shown in Figure 12, the low risk group selected faces at chance levels (χ^2=3.024, 4 d.f., ns).

A Mann-Whitney test indicated that women in the high conception risk group were significantly more likely to choose a masculine face than those in the low conception risk group (U=1798, Z=-2.255, p=0.024, Figure 12). This indicates a cyclic preference for male face shape.

If subjects responded with equal likelihood across the menstrual cycle, we would expect approximately 45 in the 'high risk' group, and 94 in the 'low risk' group from the number of cycle days represented by each group. The 'high risk' group, with 55 members, was slightly but not significantly over-represented (χ^2=3.513, 1 d.f., p=0.061).

A separate analysis of subjects using oral contraception (n=39), using the same time periods for classification, did not reveal a cyclic preference (U=127, Z=-1.28, ns) although sample size was low ('high risk' n=13, 'low risk' n=26). Participant age had no significant correlation with attractiveness judgements for neither the pill and non-pill groups combined, (r_{rho}= 0.043, n=179, ns) nor the separate groups (high risk, r_{rho}= 0.004, n=55, ns; low risk r_{rho}= 0.063, n=84, ns).

5.3.4 DISCUSSION

These preliminary data indicate that women are attracted to exaggerated male traits when conception following coitus is most likely (the follicular phase), and not at other times of the menstrual cycle. The absence of a marked preference for any of the faces in the low conception risk group (menses and luteal phases) is consistent with Gangestad & Thornhill's (1998, in press) findings that women discriminate between the scent of symmetrical and less symmetrical men during the follicular phase, but not at other times.

Despite the apparent preference for feminised faces in Figure 12, no significant overall preference for 'feminised' male faces was found in this study, contrary to previous work (Perrett et al, 1998; chapters 2-4). There are several possible reasons for this failure to replicate. Although the difference just failed to reach significance, more subjects responded in the follicular phase than would be expected. Women may have been more likely to respond to a study on male attractiveness in the follicular phase than other phases due to an increase in sexual interest, or possibly because it is easier to remember the date of onset of the previous menses when the event is relatively recent. As women who were in the follicular phase were more likely to prefer a masculinised face, the overall preference would be shifted slightly toward masculinity due to this over-representation. Secondly, unlike previous studies reported by Perrett et al (1994, 1998 experiments 1-3 above), the age of the participants was not linked to the age of the faces that comprised the composite. As 'masculinised' faces are rated as older than 'feminised' faces, the older women in the current sample may have been slightly biased toward masculinity. This was not apparent from the data – age did not correlate with the masculinity of faces chosen. Finally, the average face employed in this study was constructed from different source faces than the faces reported in the previous study.

Other methodological points suggest caution before drawing firm conclusions from these preliminary data. The reliability of self-reports of menstrual cycle phase are disputed (Bean et al 1979, but see Baker et al, 1998), and of course, not all women have a 28 day menstrual cycle. A study employing hormonal assays to ascertain cycle phase would be an improvement over this study but it is difficult to see how the methodology employed here could introduce a systematic bias to the results rather than noise. In addition, self-selected questionnaire respondents may not be representative of the population at large, although it seems unlikely that this alone could account for the findings. It is also possible that the interpretation of the task may have changed across the menstrual cycle. Subjects were asked to choose the 'most attractive' of the five faces presented: criteria of 'attractiveness' may change across the menstrual cycle with fluctuating sexual interest.

Despite these drawbacks, experiment 3 has found evidence of a cyclic preference for male face shape and that women exhibit preferences for biologically relevant aspects of facial structure that may signal heritable genetic characteristics. The results are also consistent with other cyclic preferences relevant to mate choice. Menstrual phase may be one variable that contributes to the difficulty of defining what females find attractive in male faces. Female sexual interest has a peak at ovulation, and preferences for men honestly advertising immunocompetence at this time may be adaptive. In spite of this, women are sexually receptive across the menstrual cycle. Sex when conception is unlikely probably serves purposes not directly linked to fertilisation (Hrdy, 1981). The current study and data from previous experiments suggest a speculative interpretation of cyclic visual preferences.

Dominance and 'quality as a parent' are attributions made to opposite ends of the continuum that relates to facial masculinity and femininity and each may be associated with behaviours that have costs and benefits to reproductive success (Perrett et al 1998; Berry & Wero, 1983). A preference for males with a more masculine (i.e. high testosterone) appearance when conception is most likely may confer benefits for offspring in terms of heritable immunocompetence but confer costs due to potentially decreased paternal investment.

In humans, paternity uncertainty results from concealed ovulation and limited visual similarity between offspring and their fathers (Christenfeld & Hill, 1995; Pagel, 1997). Rates of extra-pair paternity are certainly non-zero, although well-controlled studies are hard to find (Macintyre & Sooman, 1991). Paternity uncertainty coupled with converging evidence that female preferences for 'male' traits peak mid-cycle suggest female mating strategy need not be entirely exclusive. As with some other species (Graves et al 1993; Andersson, 1994), human females may have been selected to pursue a mixed mating strategy under some ecological and social conditions. Women with a main sexual partner are more likely to have extra-pair copulations in the follicular phase of the cycle than during the luteal phase or menses (Baker & Bellis, 1995). Some females may choose a primary partner whose relatively low masculine appearance suggests cooperation in parental care, whilst occasionally pursuing extra-pair copulations with males with a relatively masculine appearance indicating good immunity when conception is most likely. Potentially, sexual behaviour arising from cyclic preferences may provide the benefits of polyandry whilst maintaining the advantages of ostensive monandry, thus realising a reproductive advantage that selection may favour.

5.4 EXPERIMENT 4

5.4.1 INTRODUCTION

Experiment 4 aimed to address some of the problems of the previous experiment. Instead of a 'single shot' methodology, a repeated-measures design was employed, that allowed testing of the same women in different phases of their menstrual cycle. This study was conducted in Japan with Japanese students, and used stimuli sets of two races (Japanese and Caucasian). Although cross-cultural consistency does not *prove* that a behaviour is an adaptation (ecological pressures may cause local adaptations), cross-cultural replication would add weight to an evolutionary interpretation of the results. In addition, at the time of the study, oral contraception was illegal in Japan. Oral contraception alters the hormonal processes of the menstrual cycle and appears to disrupt cyclic and other preferences (e.g. Thornhill & Gangestad, in press; Wedekind et al 1995). The young Japanese women tested here are very unlikely to have used hormonal contraception and hence provide a good population with which to test hypotheses based on normal cyclic changes in homonal activity.

If cyclic preferences are an adaptation to maximise fitness in offspring through extra-pair copulations, it is possible that women within relationships may manifest different patterns of response than those without a partner. Previous research suggests that women in all cultures value positive personality characteristics in a partner above all else (Buss, 1989). Recent work has indicated that stereotypical personality characteristics do indeed influence attractive judgements (Perrett et al, 1998; expt. 1 above). Once women have secured a partner, however, those seeking extra-pair partners may have considerations other than personality foremost in their mind. For example, cues to possibly heritable quality (as signalled by androgenic effects on face shape) may become more important. This experiment investigated such a possibility.

METHOD

SUBJECTS

Japanese subjects (n=75, mean age 21) answered questions giving average cycle length and date of previous menses onset. Thirteen subjects reported irregular menstrual cycles and were excluded from further analyses. From this information, two testing sessions were arranged allowing sampling of individuals' preferences during both high and low conception risk phases of their menstrual cycles. For women who reported a cycle length of greater than or less than 28 days, ovulation date was estimated at 14 days prior to menses onset using a backwards counting methodology to (Regan, 1995; Urdy & Morris, 1977). Subjects informed the experimenter of the date of the onset of their next menses by post to check that the sessions had been arranged successfully in the desired 'windows'. Sessions after ovulation and before menses onset (the luteal phase) were classified as 'low conception risk'; sex in these phases is unlikely to result in pregnancy. Sessions in the follicular phase (between the end of menses and ovulation) were classified as 'high conception risk'. Thirty-nine subjects were successfully tested in both 'low' and 'high' conception risk phases. Thirty-six subjects were only tested once, or presented details that precluded any certainty of cycle phase (i.e. extremely long cycles in excess of 40 days). Data from these subjects were not analysed further. No subjects used oral contraception (which was illegal in Japan at the time that the study was performed).

STIMULI AND PROCEDURE

In both testing sessions, subjects were asked to select the face they considered most 'physically attractive' ('miryoku-teki') from five Caucasian and, separately, five Japanese male faces (40% and 20% feminised, average and 20% and 40% masculinised: manipulated with sexual size dimorphism, section 7.3.2). Subjects were also asked if they currently had a 'steady boyfriend' ('tsukiatteiru'). Stimuli were symmetrical versions (section 7.4) of those used in Perrett et al 1998 (set 1: KJL, and set 2: KJL_JAP, sections 7.1.4.1 & 7.1.4.2) but with no masking of hair, presented with full counterbalancing (figure 13). The experimenter presenting the stimuli was a female Japanese student (R. Minnasawa) who was blind to the experimental hypothesis.

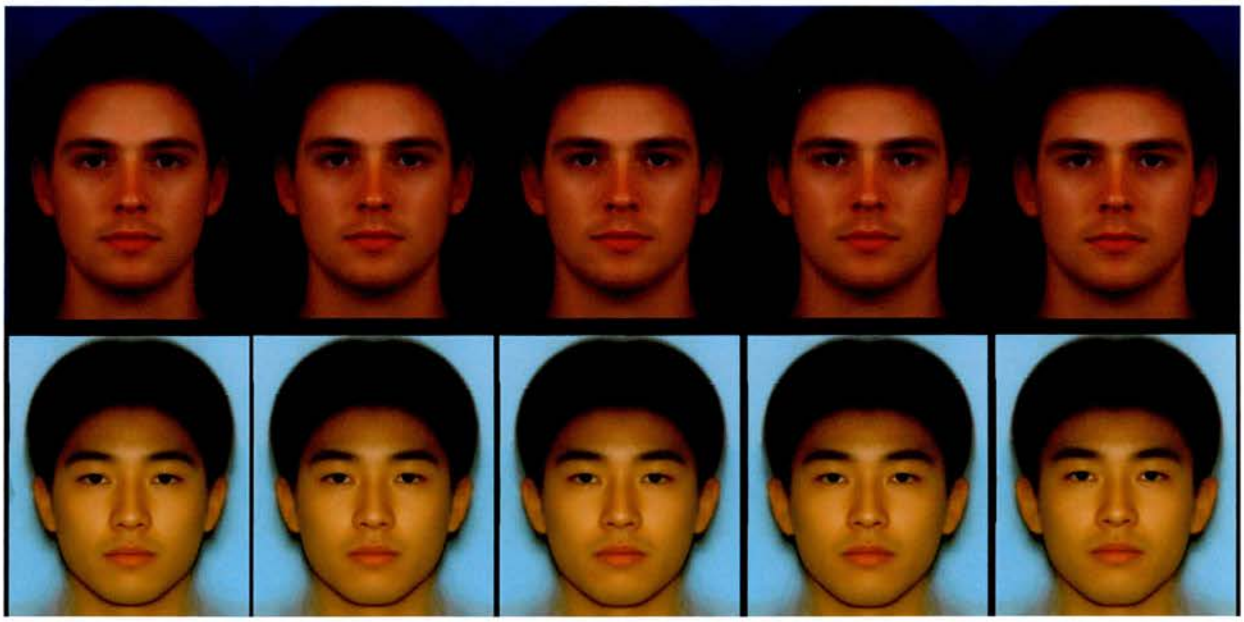
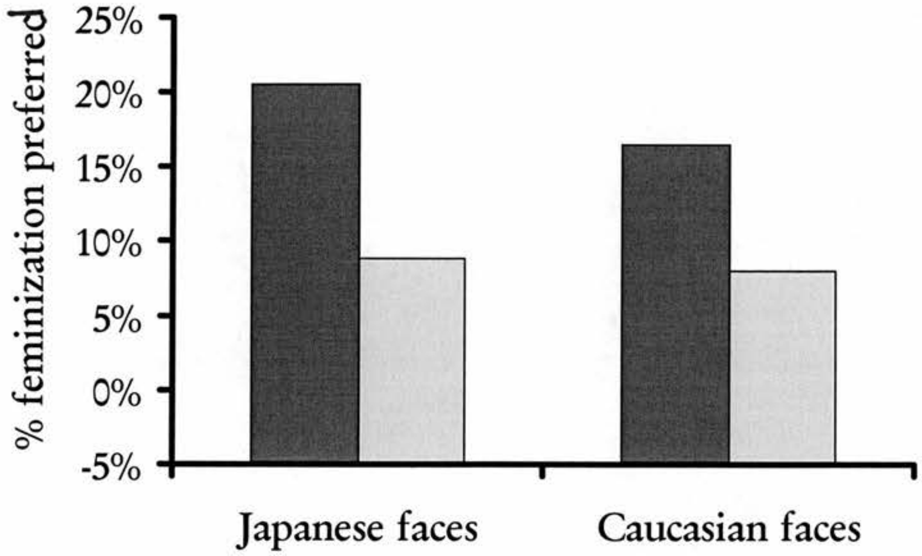


Figure 13: Caucasian (top) and Japansese (bottom) stimuli used in experiment 5. From l-r: 40%, 20% feminised, average, 20%, & 40% masculinised.

5.4.2 RESULTS

A 2*2*2 repeated measures ANOVA with factors of stimuli race (Caucasian and Japanese), conception risk (High and Low) and relationship (subject had or did not have a partner) was used to analyse the data.

The analysis showed a significant main effect of conception risk ($F_{(1,37)}=9.47$, $p<0.004$), with subjects preferring less feminised faces in the high conception risk phase than in the low conception risk phase. There were trends that women with a partner preferred more masculine faces ($F_{(1,37)}=3.59$, $p=0.066$) and underwent a greater cyclic change in preference than those without ($F_{(1,37)}=3.2$, $p=0.08$). No effects of stimuli origin (Caucasian or Japanese, $F_{(1,37)}=0.4$, $p>0.5$), or any interactions were found (race*risk, race*partner, race*risk*partner: all $F_{(1,37)}<14$, all $p>0.25$). These findings are summarised graphically in Figure 14



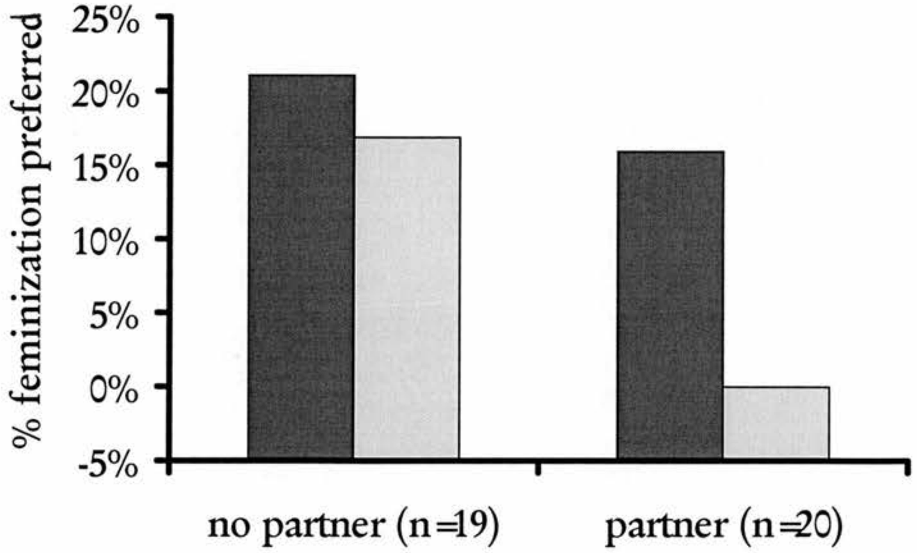


Figure 14: Effects of conception risk on femininity preferred in Japanese and Caucasian faces (top) and for subjects with (n=20) and without (n=19) a partner (below).

5.4.3 DISCUSSION

Experiment 4 adds usefully to the data of experiment 3 in several ways. Firstly, as in experiment 3, subjects have demonstrated a significant shift toward more masculine male face shapes during the follicular phase of the menstrual cycle. The methodology of this experiment represents a significant improvement over that of the survey study, as it employed a within subjects design with full counterbalancing of stimuli. Additionally, two stimuli sets (one Caucasian and one Japanese) were used. The question asked ('physically attractive') is less open to misinterpretation than the question used in Expt. 3 above.

This experiment investigated the effects of a current relationship on cyclic preferences. Whilst no significant effects or interactions were found, there were trends that women in relationships preferred more masculine faces in general and showed a greater follicular phase shift towards masculinity than women without partners. Both of these trends are consistent with an extra-pair copulation interpretation of cyclic preferences: once a main partner has been acquired, women may be less concerned with perceived personality indicating paternal behaviour, and more concerned with cues to genetic quality in possible EPC partners (i.e. those features associated with testosterone). These cues may be especially important when EPCs are most likely to occur and when conception is most likely during the follicular phase of the menstrual cycle.

5.5 EXPERIMENT 5

5.5.1 INTRODUCTION

Experiment 5 expands upon the findings of experiments 3 and 4 in a number of ways. In common with experiment 4, this study is laboratory based and within-subjects, using repeated testing of UK University students across the menstrual cycle. An interactive methodology is used allowing subjects to alter the apparent masculinity and femininity of the faces presented with more sensitivity. To further investigate the possible role of cyclic changes in low paternal investment or extra-pair sex, subjects were asked to make one of two attractiveness judgements: to pick a face that would be attractive in a long-term partner, or a face that would be attractive in a short term-partner.

5.5.2 METHOD

STIMULI

Five facial continua (Figure 15. Sections 7.1.4.1-5) were prepared from separate, unmasked symmetrical composites. 50% feminised and 50% masculinised versions of these stimuli were prepared with size dimorphism, and interactive continua were generated allowing subjects to manipulate the between these faces in real-time on PC, by moving a mouse pointer left or right on screen. Each continuum consisted of 11 images ranging from 50% feminised to 50% masculinised in 10% steps. Mouse movement caused an apparently smooth transition between each image, by replacing the image displayed with one adjacent to it in the continuum. Order and left-right presentation were randomised.

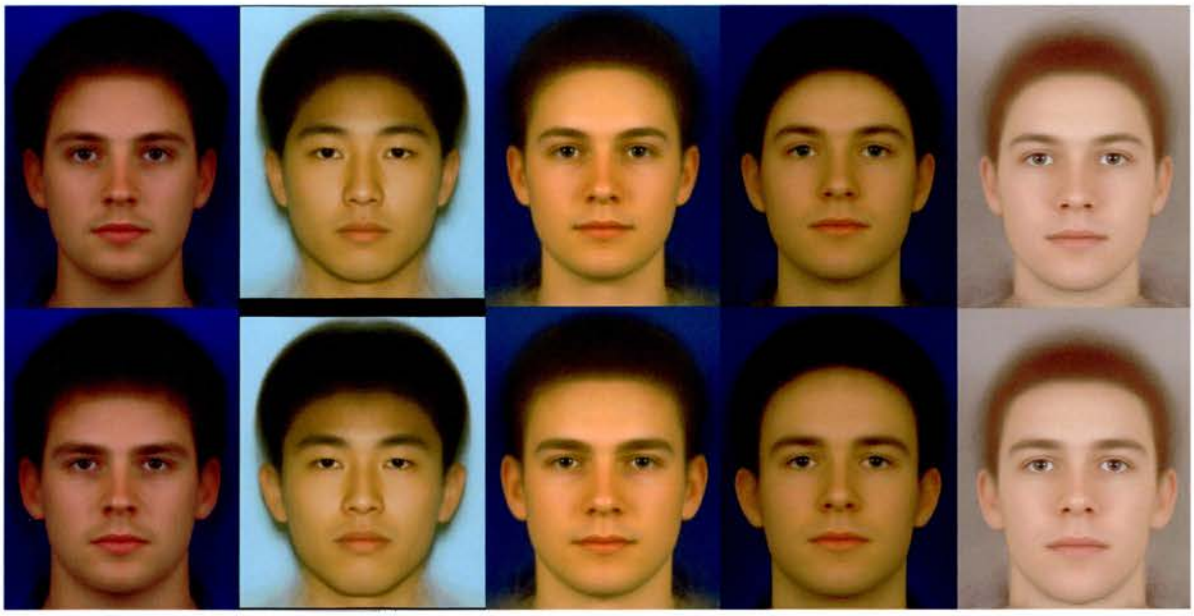


Figure 15: 50% feminised (top) and 50% masculinised (bottom) stimuli from the five continua used in experiment 6. From l-r: KJL; KJL_JAP; TW; IPV2; JEN

SUBJECTS AND PROCEDURE

From an original sample of approximately 85 British participants, 65 subjects (mean age, 20 years), completed three or four sessions at approximately weekly intervals, and gave details of their menstrual cycle and use of oral contraception. Each session was classified as either 'high' or 'low' conception risk. As the accuracy of self-reported menstrual status is questioned (see above) we assumed that the session nearest menses onset would represent the most accurate recollection, and so phases were estimated from this date. Where possible, a retrospective model of the menstrual cycle was employed, with ovulation occurring 14 days before menses onset. Retrospective models of the menstrual cycle are claimed to be the most accurate due to the fixed lifespan limitations of the corpus luteum (Chabbert Buffet et al, 1998; Regan, 1996). Thirty of the subjects returned 'cycle update' forms giving the date of onset of their next menses after the end of the testing period allowing more accurate determination of cycle phase.

Where a retrospective model of the menstrual cycle was inappropriate (in subjects whose period corresponded with the first week of testing, who then completed their experimental sessions before next menses onset) ovulation was assumed to have occurred on day 14 of the cycle.

Subjects were asked specifically to choose from the continua the most attractive face for a 'long-term relationship' or a 'short-term sexual relationship'. 27 made long-term and 28 made short-term judgements in all sessions. A further 10 subjects made both judgements. Responses to each face from each subject were averaged over 'low risk' and over 'high risk' sessions.

5.5.3 RESULTS

Repeated measures ANOVA for subjects not using oral contraception, ('short-term' n=17; 'long-term', n=20; both judgements, n=6) showed no main effect of conception risk ($F_{(1,47)}=2.13$, $p=0.151$) or type of relationship ($F_{(1,47)}=0.017$, ns). Conception risk, however, interacted with type of relationship (short- or long-term; $F_{(1,47)}=5.39$, $p=0.025$). When judging attractiveness for a short-term sexual relationship preferred face shape was less feminine during the high conception risk phase. In contrast, preferences remained constant when women judged attractiveness for a long-term relationship (Figure 16). Subjects preferred different levels of femininity in different composites ($F_{(4,188)}=11.1$, $p<0.01$; figure 17), but the lack of significant interaction (risk*stimuli*relationship type $F_{(4,188)}=0.79$, $p=0.53$) showed that the cyclic change in preference was present for all stimuli. Analysis of data from subjects using oral contraception (n=22, 4 subjects in both long and short conditions) revealed no cyclic changes in face shape preference ($F_{(4,96)}=0.001$, $p=0.97$) or interactions.

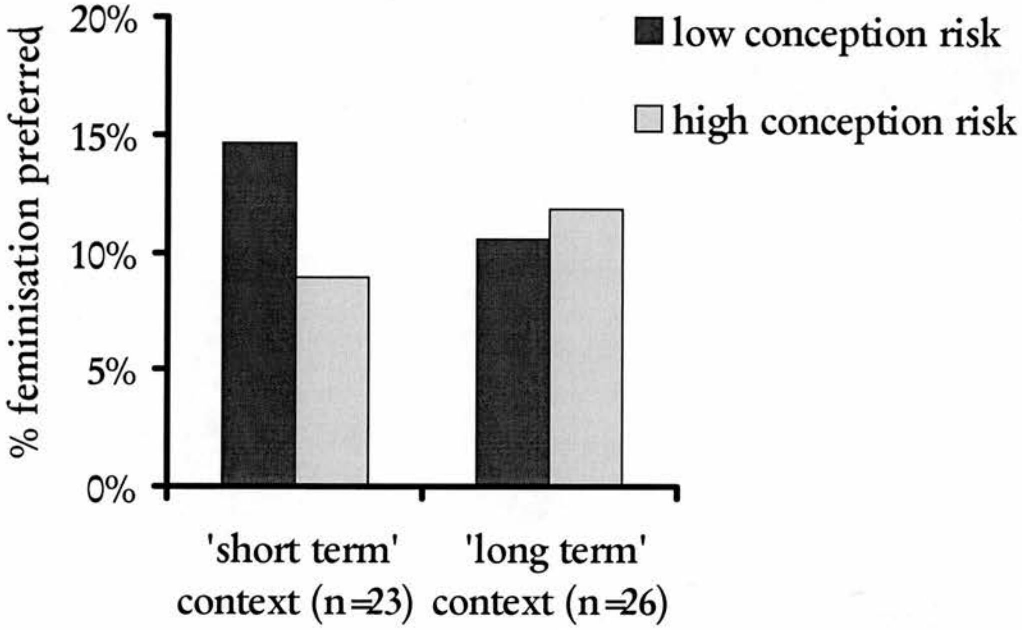


Figure 16: UK longitudinal study data. Effects of relationship context and conception risk

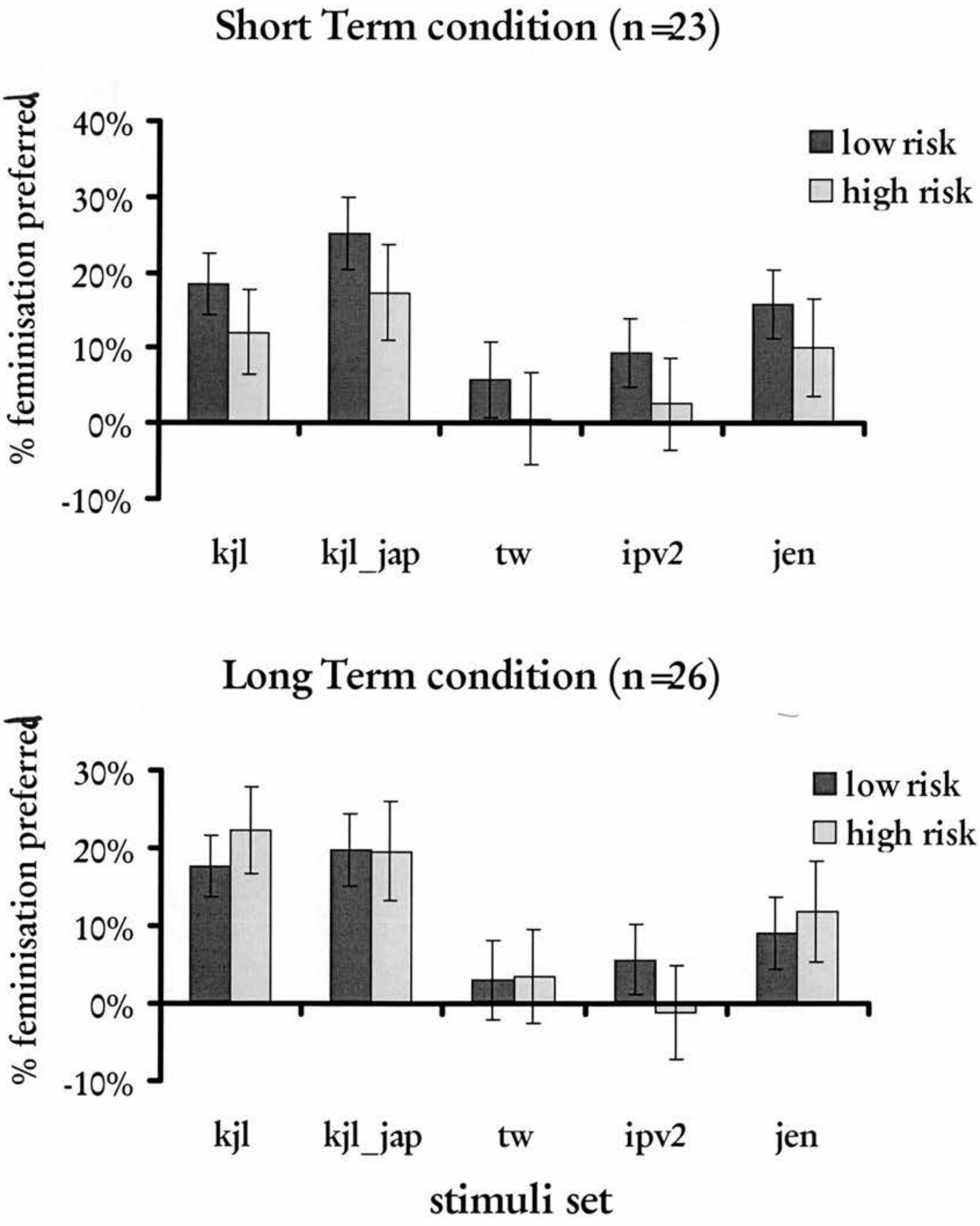


Figure 17: responses to each individual face set in each condition at high and low risk

5.5.4 DISCUSSION

Again, the data from experiment 5 supports the hypothesis that cyclic changes in female preferences favour more masculine male face shapes when conception is most likely. However, the experiment 5 adds important new data: these changes do not occur when females choose faces to be attractive in a long-term partner. When judging attractiveness in a short-term sexual partner, preferences favour more masculine looking faces. These findings are consistent with an extra-pair copulation interpretation of cyclic preferences.

Females seeking a long-term partner in this experiment always prefer the same level of feminisation in male face shapes across the menstrual cycle; such shapes elicit favourable personality attributions that may be related to behaviour (chapter 3 above). When likely paternal investment is low however, as in short-term sexual encounters such as EPCs, more weight is given to characteristics that may indicate 'good genes' in partners, through facial traits thought to be linked to androgens. To some extent, this finding reconciles the overall preference for femininity in male faces with other literature on male sexual activity. As reported earlier, dominant (presumably high testosterone) men copulate earlier (Mazur et al, 1994) and achieve more sexual access as adults (Perusse, 1993). Townsend et al (1995) report that high status male college students such as athletes reported often more than 100 sex partners. Such a number of partners obviously reflects low male investment in each female, and vice-versa. It has been suggested that this low investment in a given women reflects the quality of the male involved: they are more attractive to women and can therefore obtain more partners, despite the women's desire for a monogamous relationship (e.g. Symons, 1979, pers. comm.; Townsend et al 1995; Gangestad & Thornhill, 1997; Wiederman & Dubois, 1998).

Nonetheless, it is widely accepted that bi-parental care is an adaptation: men and women desire a long-term pair bond as it has led in the past to greater reproductive success. Men do not have to be 'tricked' into socially monogamous relationships: strategies of no male parental investment have been selected against. High testosterone men apparently fail to either attempt or succeed in marital relationships (Mazur & Booth, 1998). To say that all women seek high testosterone men as long-term partners on the basis of the frequency of college students' short-term sexual relationships may be an oversimplification of a complex set of context dependent female mating decisions. The data of experiment 5 imply that higher testosterone men may be favoured for short-term sexual relationships when conception is likely, whilst men with lower testosterone, with their concomitant likelihood of extended paternal investment, are preferred for long term relationships.

5.5.4.1 Differences between continua: what is averageness?

Although the pattern of cyclic responses was the same across all the continua, it is clear that participants in this experiment prefer different levels of femininity in different composites (Figure 17 above). An overall preference for femininity was found in the composites used in Perrett et al, 1998 (KJL and KJL_JAP) and in a third set (JEN). In the remaining two face sets, TW and IPV2, the preference for femininity was marginal. Set TW was also used in experiments 2 and 3: there was no clear preference for femininity in male faces in either of these experiments.

One possible factor contributing to these discrepancies is the age of the individuals who made up the face set. The source faces in KJL and KJL_JAP are indeed slightly older than those in TW and IPV2. Set JEN, however, has the youngest constituent faces and yet a clear preference for femininity is found. Post hoc explanations to account for the differences in preferences between different stimuli can be generated (e.g. different photographic conditions between sets; different exclusion criteria in the amount of facial hair tolerated from set to set) but the explanatory power of such hypotheses is limited.

The differences in responses to a variety of composite stimuli focus serious questions about testing of the facial averageness hypothesis. Proponents of the averageness hypothesis suggest people are looking for faces which represent the average of all the exemplar faces that they have been exposed to, as such faces may indicate high levels of heterozygosity. But how, for example, do such mental prototypes represent age? Cartiliginous tissue continues to grow throughout life; so, youthful faces have smaller than average noses and ears - and yet are consistently found more attractive. The presence of cyclic shifts in preferences, however, renders such questions less important than they might otherwise have been. Women *cannot* prefer average faces if their preference for faces changes cyclically.

5.6 General summary of chapters 1-5

On the evidence of the experiments 1-5, female preferences for even static images of male faces represent a complex set of decision-making processes, summarised below.

Perrett et al (1998) indicates that women do not have clear preferences for 'masculinised' (high testosterone) face shapes as predicted by indicator models of sexual selection. Experiment 2 indicates that this is not due to the stimulus artifacts of masking or the loss of size dimorphism.

Experiment 1 indicates that stereotypical personality judgements attributed to static faces appear to influence attractiveness judgements. Masculinised faces (indicating high levels of androgens) are considered to posses fewer desirable personality traits than feminised faces. These attributions may have some validity, and a reasonable (although as yet unsupported) biological model linking androgen levels, behaviour and facial shape fits in with the observed preference pattern and, apparently, the fossil record: an overall preference for relatively low testosterone men may be an adaptation.

Biological facial characteristics that are considered putative indicators of 'good genes' are not, however, ignored. They appear to be appraised in the light of stereotypical personality judgments and in the context of life history factors (such as the type of relationship sought and possibly the relationship status of the woman). Furthermore, preferences are mediated and interact with cyclic hormonal changes linked to the likelihood of conception following sex. Relative masculinity seems to be preferred at times when conception is most likely.

These preferences have been found in two cultures (UK and Japan), with student aged (expts. 4 & 5) and older women (expt. 3), and with multiple stimuli sets. If preference data is reliably linked to actual sexual behaviour (a question that should be addressed by future research), a model linking likelihood of parental investment and facial masculinity preferred by females is suggested: when parental investment is sought (i.e. for a long-term relationship) facial shapes associated with relatively lower testosterone levels are preferred. Such faces may reliably indicate pro-social personality characteristics. When, however, likelihood of parental investment is low (in short-term relationships or possibly extra-pair copulations when the likelihood of conception is high) relatively more masculine faces are preferred, in a fashion more consistent with 'good genes' hypothesis.

The results of the studies reported here have some caveats, however, that can be addressed by future research. Firstly, a replication of an overall preference for femininity and cyclic shifts in such preferences using real, rather than composite, faces would add strength to both arguments. Facial-metric measurements of such stimuli may also isolate the features driving attractiveness judgements and allow Cunningham's multiple motives model to be tested thoroughly.

Although the preferences described so far in this thesis generalize to participants in the UK and Japan, further cross-cultural work is needed to investigate the possible consequences of local environmental conditions. Yu & Shepard (1998) indicate that preferences for low female waist-to-hip ratios (e.g. Singh, 1993), are not universal, as previously thought. Similarly, preferences for feminised male faces may not be universal: in societies living under the pressure of a higher parasite load than modern UK and Japan, possible indicators of good genes (i.e. testosteronised features)may be given more weight in attractiveness judgements.

6 Facial Similarity and Judgements of Attractiveness

Anecdotally, spouses are often said to resemble one another. Despite overall agreement on 'attractive' and 'unattractive' faces between and within populations, individual differences in preferences certainly exist. Chapters 1-5 reviewed the effects of manipulating facial sexual dimorphisms on attractiveness judgements. As summarised above, female preferences have been shown to vary according to a variety of factors such as personal situation and menstrual cycle phase. Another possible influence on an individual's preference is self-similarity: are people attracted to individuals who look similar to themselves? This chapter reviews evidence for such preferences, and reviews the work of researchers who suggest that similar partners may have a biological advantage over dissimilar couples. Finally, two experiments that use computer graphic manipulations of faces that alter the similarity between the stimuli and subject are presented.

6.1 An overview of assortative mating in man.

Propinquity, physical appearance, health and social status are among the many components that influence human mate choice, and ensure that *panmixia* (random mating) does not occur in humans. The most widely reported mating pattern in vertebrates appears to be *assortative* mating (Burley 1983); the formation of pairs between individuals who are more similar to each other on a given trait than would be expected by chance (Thiessen & Gregg, 1980). This section examines the hypothesis that 'like prefers like' concentrating on definitions of assortative mating and reviewing the evidence for assortative mating in humans. Section 6.2 will discuss the possible advantages and disadvantages of assortative mating patterns from an evolutionary perspective, with respect to the individual and the population. The concept of some optimal amount of assortment will also be introduced here.

6.1.1 Definitions of, and evidence for assortative mating.

Mating is described as assortative when individuals with certain phenotypes mate at levels above chance (Partridge 1983). Positive assortative mating (also called *homogamy*, or more simply just assortative mating) occurs when the members of a pair are phenotypically more similar than would be expected if mating was random. Negative assortative mating (*heterogamy* or disassortative mating) occurs between individuals whose phenotypes are more dissimilar than chance predicts (Burley, 1983). Inbreeding is subtly different from assortative mating, yet the concepts are often confounded. Inbreeding (or consanguinity in humans) involves relatedness between the partners, and is often considered a special form of positive assortment (Vandenberg 1972). However, outside of human studies (in which inbreeding & incest are culturally defined) most researchers are unwilling to state the level of relatedness at which inbreeding becomes assortative mating or vice-versa, and I will use the terms loosely, with inbreeding referring to extreme assortment. Data accumulated from many species indicate that positive assortment is the most common pattern found among animals (Burley, 1983, Thiessen & Gregg, 1980). In humans similar evidence is found, which will be briefly discussed below.

6.1.2 Measurement of anthropometric characteristics.

Early research into human assortative mating focused on large scale studies in which anthropometric characteristics were measured, and correlated between spouses. Spuhler, 1968, reports two such studies that he carried out. His study of partnership in Ann Arbor, Michigan, involved taking 43 physical measurements (such as arm length, ear lobe length, hand breadth etc.) from 205 married couples. He found significant positive correlations on 29 of these measurements. Interestingly, 7 of 15 measured facial or cranial correlations proved significant. No negative correlations were reported. Spuhler's second study measured the physical characteristics of 78 married couples from a population of Ramah Navaho native Americans in New Mexico. None of the 40 measurements showed a significant correlation, indicating that assortative mating for physical characteristics is not a ubiquitous finding in humans.

In addition to his own studies Spuhler presents a comprehensive review of the early studies of partner similarity. He collected correlation coefficients between 105 physical characteristics from research spanning nearly seventy years (From Pearson's 1903 study of eye colour in Britain to his own Ann Arbor study). The overall pattern is one of low positive correlations (0.01- 0.35) on many physical features - all but 41 of the 288 correlations reported are positive, and Spuhler only reports two of the negative correlations as significant, compared to 38 of the positive correlations. A similar review was carried out by Roberts (1977), with similar conclusions; weak positive correlations between partners appear to be the most frequent result.

Although a far more extensive review of the literature would be possible, the value of repeating the work of Roberts (1977) and Spuhler (1968) here is questionable. Many of the early studies employ methodology that precludes firm conclusions being drawn. Few take into account that age may have an influence on partner similarity; lip thickness is known to decrease with age (Rowland 1995), for example. The effects of environmental coexistence on the correlations are also largely ignored - the act of living together and sharing similar diets may have serious consequences on measures such as abdominal circumference. The fact that many of the measures are intercorrelated is stressed by Spuhler. For example, it seems unlikely that upper arm length will increase without an accompanying change in forearm length, so the large number of correlations may give a false impression of the number of traits that are assorted for. Recent studies of anthropometric similarity are better controlled. A recent study of assortative mating for body weight found a significant correlation between spouses that could not be explained by cohabitation (the 500+ couples involved were weighed both immediately prior to marriage and twenty years later) or age (Allison 1996). Malina et al's (1983) study of assortative mating in a Mexican community also partialled out age and found evidence for assortative mating for a number of characteristics.

Large scale studies of spousal anthropometric characteristics have been infrequent since the late sixties, possibly due to their expensive and time consuming nature, coupled with a lack of new or surprising findings. Nevertheless, it seems reasonable to conclude from the sheer volume of the early and the precision of the later studies that some physical homogamy occurs in human marriage, and not all of this partner similarity can be accounted for by age or coexistence effects.

6.1.3 Ratings of perceived couple similarity.

Some studies have directly studied the perceived similarity of members of real married couples, usually by presenting subjects with a selection of images and asking them to match up pairs. Griffiths & Kunz, (1973), took photos of actual married couples. Each member of the couple was photographed individually, and a common background was used for all 30 of the couples used in the study. The couples were split into six groups by the length of the marriage. The test involved presenting each group of ten faces to 295 subjects, who were asked to identify the couples. Within group ages and socioeconomic status were homogenous, to prevent subjects using cues such as clothing or hair styles. The results of the study are intriguing: although couples married for less than ten years were matched at levels above chance, subjects failed to successfully match couples married for between ten and twenty years. Curiously, couples married for over 20 years were once again successfully identified by the subjects. This is obviously a difficult result to interpret; similar people appear to be attracted to each other, and then grow dissimilar in the middle years of their marriage before resembling each other again in their middle age and later life. However, the small stimuli sets used (just five couples in each group) may go some way to explaining the curious patterns in the data.

Zajonc et al performed a similar experiment to test the hypothesis that couples do not get together due to similarity, but become more physically alike over time due to sharing similar diets, lifestyles and emotional experiences (Zajonc, Adelmann, Murphy, & Niendenthal, 1987). To test this theory of environmental coexistence, Zajonc et al obtained two photographs from each individual in twelve married couples. One of these photos came from the first year of the marriage, and the second was taken in the twenty-fifth year of the partnership. The photos were cropped and masked to reveal only the head and face of each person, and then split into 'young' and 'old' subsets. None of the subjects were exposed to the young and old versions of the same face. Target faces were presented with six opposite sex test faces, one of which was the target's real life partner. The subject's task was to rank either the similarity or the likelihood of marriage between the test faces and the target, depending on the condition. Zajonc et al's results indicated that for the 'young' subset subjects did not perceive a level of similarity (or a likelihood of marriage) between the target and partner any greater than would be expected by chance. Conversely, real couples from the 'old' subset were ranked as more similar and more likely to be married than chance predicts - a result that supported the environmental co-existence hypothesis. The Zajonc study is notable for a comprehensive design, using sophisticated counterbalancing and an interesting control condition in which the faces of the stimuli were obscured. Nevertheless, sourcing the stimuli photographs from the subjects' own collections may raise some doubts as to the validity of the results. In addition, the statistical analysis appears slightly inappropriate, using multiple t-tests on seemingly non-parametric data. These results are somewhat contrary to the Griffiths & Kunz (1972) study. Although both studies indicate some homogeneity between couples married for twenty five years or more, the Zajonc et al study does not demonstrate any similarity between newly wed couples. An interesting observation that Zajonc does not stress is the lack of any difference between the 'similarity' and 'likelihood of marriage' rankings. This indicates that Zajonc's subjects associated facial resemblance with the likelihood of marriage between individuals; an example of the common idea that people marry people who look like themselves.

Hinsz (1989) conducted a third study of facial similarity within real life partnerships. The stimuli in this experiment were photographs of the individuals in 30 engaged couples and 30 couples who had been married for around 25 years. The subjects were presented with opposite sex pairs of photos, and then asked to rate the similarity between the two faces. Half of the pairs presented were actual couples, and the other half were randomly generated couples. Hinsz used both 'young' (undergraduate) and 'old' raters (recruited from 25th wedding anniversary adverts). A significant real v random couple effect was discovered; real couples were rated as more similar than randomly generated couples. Unlike the Zajonc et al study, couples that had been together for longer periods of time were not perceived as more similar than new couples in the Hinsz experiment; thus no support for the environmental co-existence hypothesis was found.

6.1.4 Assortative mating for psychological traits and cultural values.

Although this review will preface a study of physical similarity between human sexual partners, it is important to note that assortative mating takes place for non-physical factors. Indeed, much stronger correlation coefficients are reported between partners for non genetically transmitted characteristics such as religion, educational level, socioeconomic status and political beliefs than for any physical characteristics (Vandenberg, 1972, Thiessen & Gregg, 1980, Rushton, 1988, 1989, Epstein & Guttman, 1982). Characteristics which are genetically heritable but with a strong environmental influence (such as intelligence) show mid-range correlation coefficients between spouses. Johnson, Ahern & Cole, (1980), reviewed studies from 1928-1973 that examine spouse similarity for intelligence. They found positive correlations between spouses between 0.12 and 0.76 across the many different test batteries employed. Bouchard and Mcgue (1981) reviewed 16 standardised IQ studies and found an overall correlation of 0.37 between partners in 3817 couples. There seems to be an inverse relationship between the genetic component of a trait and the amount of assortment that occurs for it (Thiessen & Gregg, 1980).

Interestingly, there is evidence that assortative mating also occurs for abnormalities. Assortative mating has been observed between sufferers of schizophrenia, alcoholism, depression and neuroses (see Thiessen and Gregg, 1980 for a review).

The evidence above demonstrates that assortative mating occurs over a wide range of characteristics in humans, with phenotypically similar individuals mating together at levels above chance. However, an assortative *pattern* is not necessarily caused by assortative (or *homotypic*) *preferences* (Burley, 1983). Assuming that 'like mates with like' because 'like prefers like' is an oversimplification; in a population where a certain characteristic is universally considered attractive (a *type* or *directional* preference) an assortative pattern can still develop. Consider a hypothetical monogamous species in which long legs were considered attractive by both sexes. Individuals with long legs would be pursued by many prospective mates, and hence would have a good selection of partners to choose from. Given the popularity of ample pins, those endowed with lengthy limbs would probably seek an equally gifted mate. Those individuals with more stunted shanks may find themselves unable to attract leggy partners despite their best efforts. They would end up with a mate of a similar status in the leg length department. An example of this type of effect can be found in Berscheid et al's classic study of mate choice in humans (Berscheid, Dion, Walster & Walster, 1971). They showed that although college students prefer to date highly attractive people in ideal situations, in situations in which rejection is possible they actually choose dates of similar attractiveness to themselves. Burley's assertion that homotypic preferences are 'often erroneously inferred' from assortative patterns, and that many studies of assortment are flawed by not studying preferences directly makes identifying type from homotype preferences problematic. The next section examines the possible benefits and disadvantages of phenotypically similar individuals mating, in an attempt to elucidate a possible theory that could lead to the occurrence of homotypic preferences.

6.2 Why positive assortment?

6.2.1 Assortment and inclusive fitness

If it is assumed that phenotypical characteristics are related to genotype (anthropometric characteristics are certainly heritable: Susanne, 1977), assortative mating must increase the inclusive fitness of the individuals involved over other mating strategies to proliferate. Several researchers have proposed both the costs and benefits of such mating patterns from this evolutionary perspective.

Many theorists propose that assortative mating is beneficial as it increases the coefficient of parent-offspring relatedness (Thiessen & Gregg 1980, Epstein & Guttman, 1982, Rushton, 1988, 1989). This coefficient is defined as the 'probability that any randomly selected allele in an individual has a copy identical by descent in the related individual' (Daly & Wilson, 1983). Simply put, it is the probability of any one of the parent's genes being represented in the progeny. For the offspring of unrelated individuals, the coefficient of relatedness, (r) = 0.5. Thiessen and Gregg (1980) reviewed human mating and claim that assortative mating was beneficial up to the point that an 'inbreeding depression' becomes evident (the 'inbreeding depression' is a reduction in fitness caused by the expression of deleterious recessive genes, discussed later). They argue that assortative mating increases the genetic relationship between the individuals involved above 0.5, and in turn, the relationship between parents and offspring increases. They propose that 'the overall effect is a gain in gene duplication without an increase in reproductive investment' (Thiessen & Gregg, 1980). Bateson (1983) cites this as an example of one of the misunderstandings of kin-selection mentioned by Dawkins, (1979), who argues that in monogamous species there is no benefit to be gained from producing extra-closely related offspring. Should siblings mate, the resultant infant has a coefficient of genetic relatedness (r) of 0.75 to the mother. However, if the sister let her male sibling mate with another female, whilst herself mating with an unrelated individual, she would gain a child of her own (r = 0.5) in addition to her brother's offspring (r=0.25). Together these two progeny would match the 'super-offspring' of the incestuous relationship. Whilst Dawkin's criticism appears logical, it could be argued that if both the sister and brother mated with unrelated, but genetically similar individuals then a 'genetic bonus' greater than either the super-related child or the two children from the siblings' unrelated partners could be realised.

Even if careful modelling shows this aspect of Thiessen and Gregg's argument to be inappropriate for human mating, it can be applied more successfully to polygynous species in which one sex competes for partners. If many males do not mate due to female choice of a dominant or particularly attractive male, a sister who mated with her brother would then introduce an infant of high relatedness in a situation in which her brother may have been unable to mate (Smith 1979). Thiessen and Gregg also argue that increased genetic relatedness 'improves the flow of altruistic behaviors' by reducing the cost of altruism to family members, employing kin selection theory (Hamilton 1964). Rushton develops the idea that assortment for genetic similarity is an important part of many aspects of mate choice. He presents evidence (based on blood type analysis) that genetic similarity increases the fecundity of human partnership (Rushton, 1988). Such a link had been proposed by Clark & Spuhler, (1959) who correlated mate similarity for 19 anthropometric variables with the number of children produced. Seventeen of these similarity- fecundity correlations were positive, but none was larger than 0.18.

Increases in marital stability are also reported in human couples who are similar in physical and psychological characteristics; Hill, Rubin & Peplau (1976) found that couples who were similar on a variety of traits were more likely to remain together than dissimilar partners. It seems possible that an increase in marital satisfaction may lead to an increase in fecundity without the need for a 'biological' increase in fertility. Many advantages of assortative mating can be explained at the phenotypic rather than genotypic level. Burley (1983) points out that similar phenotypes may be behaviourally compatible, increasing the viability of the couple.

Partridge (1983) proposes that in many species members of the local population tend to be more related to each other than members of the wider population. Within population mating results in an assortative mating pattern that is beneficial as it avoids the costs of leaving the immediate environment to find a less similar mate. In addition, it may maintain adaptation to the local environment in offspring. She cites evidence from studies of the feeding habits of oystercatchers and ground finch, in which assortative mating is found for beak size, which is known to be related to diet. Bateson (1983) expands on these ideas and adds that different skills acquired by animals raised in particular habitats could 'disrupt' parenting, reducing the fitness of offspring.

6.2.2 The inbreeding depression.

The greatest limiting factor on a hypothesised increase in fitness caused by assortative mating is the *inbreeding depression*. Partridge (1983) claims that many natural populations do hold deleterious recessive alleles. If an individual is homozygous for these recessive genes (i.e. has two identical copies of a gene at the same locus), fitness would be lowered. The deleterious consequences of excessive homozygosity have been observed in many non-human species (see Partridge, 1983, for a review). In humans, Seemanova (1971) studied the offspring of 141 Czech women who had 161 children. All of the children were fathered by a member of the immediate family of the mother (fathers, brothers or sons). 15 of the 161 children failed to reach one year of age, or were stillborn. Five out of a non-inbred control group of 95 children perished under similar conditions. A far more striking statistic is that nearly 40% of the consanguineous group suffered severe developmental disorders. Physical defects such as heart abnormalities, deafness, and dwarfism were reported. Just over 5% of the control group suffered similar afflictions.

The inbreeding depression is the most dramatic consequence of very closely related sexual partners. However in evolutionary terms it is unstable and relatively short-lived, as deleterious recessive alleles are expressed, selected against, and eliminated from the population. Other costs of inbreeding are less spectacular yet longer lasting. The increased fixation of alleles that results from assortative mating leads to homozygosity (when both alleles at a locus are identical) in the offspring. Homozygosity can incur costs in fitness. When both forms of an allele are present at the same locus (heterozygosity), the dominant and recessive alleles may interact in an advantageous fashion (Maynard Smith, 1978). This can occur even if the recessive allele is deleterious in its homozygous form; for example individuals who are heterozygous for the sickle cell allele gain both an increase in fertility and some immunity to malaria (Raven, 1988). An increase in homozygosity leads to a lack of genetic recombination and variability in the population, which in changing environmental circumstances may reduce fitness. Maintaining heterozygosity (and therefore genetic variance) may allow at least some offspring to survive in difficult conditions. This lack of variation may also cause increased competition between similar siblings even under stable environmental conditions (Bateson, 1983).

6.2.3 Parasite driven sexual selection, heterozygosity and negative assortative mating for MHC.

The role of heterozygosity and parasite driven sexual selection should not be overlooked. Biologists have argued that conflict between parasites and hosts is the principal cause of sexual reproduction throughout nature (Hamilton, 1980, Tooby, 1982, Zuk, 1992), and the possible source of varying selection pressures that makes indicator models of sexual selection possible (section 1.3.3 above). Sexual reproduction in species with relatively long generation times provides a method of combating the threat of rapidly evolving (short generation time) parasites: epidemics are the biggest causes of mass fatality throughout human history (Ridley, 1993). Thornhill and Gangestad (1993) argue that heterozygous individuals are more likely to carry rare alleles, and therefore posses more passive genetic resistance to pathogen infestation. This could lead to selection for individuals who seek to maximise heterozygosity in offspring by outbreeding with genetically different partners i.e. selection for negative assortative mating.

One interesting and pertinent example of negative assortative mating in humans has recently been published. Human females are able to detect and respond preferably to body odours that indicate the presence of a different major histocompatability complex to their own (Wedekind, Seebeck, Bettens, & Paepke, 1995) . Wedekind et al's experiment involved asking female students to rate the pleasantness of body odours on t-shirts that had been worn by male volunteers in controlled circumstances. Humans appear able to detect differences in MHC from odour cues. The MHC antigens of both male and female subjects were typed, and it was found that women preferred the odors of men whose antigens were on average dissimilar to themselves. Wedekind et al propose that this finding has implications for the parasite-host arms race, allowing humans to react to pathogen pressures. Given the role of the immune system in facial development and attractiveness (e.g. Thornhill & Gangestad, 1993 & chapters 1 & 2) this finding should not be overlooked in the current study. Parasite pressures could lead to selection for negative assortative mating, to increase heterozygosity in the population. Another experiment by Ober et al (1998) demonstrates that fetuses that share HLA alleles with their mothers are less likely to survive to full term, providing evidence that assortative mating may have fitness costs: negative assortative mating for HLA alleles may reduce rates of spontaneous abortion.

6.2.4 Optimal outbreeding.

It can be seen that inbreeding and outbreeding have both costs and benefits, which depend on the mating strategies of the species involved, the habitat and other factors. Individuals who showed a preference for mates with whom they shared a certain amount of genetic material (i.e. a certain amount of relatedness) could optimise the costs and benefits of assortative mating. The ensuing increase in fitness would mean that this behaviour spread through the population. Although the idea that some kind of balance between in- and outbreeding is favorable has existed for some time (Wright 1933), the best experimental work in the area has been carried out by Bateson. Bateson conducted a series of mate choice experiments with Japanese quail (Bateson 1978, 1980, 1982,1983, & 1988). Stimuli birds were isolated from each other in individual compartments that were fitted with one way screens. This allowed the test bird to move around the apparatus looking in at the stimuli birds, but without any interaction. Preferences were measured by recording the amount of time each test bird spent in front of the stimuli birds. The results of the experiments shed light on the possible proximate mechanisms of mate choice and preference development, which I will return to in section 6.3. Here though, it is sufficient to note that test birds of both sexes preferred to spend time near to first cousins rather than siblings, or unrelated individuals (Bateson, 1982). Bateson hypothesised that birds were choosing partners that were 'optimally discrepant' and therefore minimised the costs of excessive in- or outbreeding. He called this effect *optimal outbreeding*. Although little other empirical work has investigated the optimal outbreeding hypothesis directly, Bateson's convincing evidence and the intuitive appeal of the idea warrant further research.

6.2.5 Genetic consequences of assortative mating in the population

If assortative mating occurs (perhaps optimally as Bateson contends) what are the effects on the genetic makeup of the population at large? Theoretically assortative mating need not have any effect on gene frequency in the population, if assortment per se causes an increase in fitness and this gain is equal for all genotypes. However, homogamy will increase the variance of the genes carried by individuals (Buss 1985). As 'like finds itself with like' more often, extreme genotypes will be found, and genes for assorted traits will increase their average homozygosity (Thiessen & Gregg, 1980; Merikangas 1982). Eventually, this may possibly influence speciation if different genotypic extremes become reproductively incompatible (Immelman 1975; Butlin 1997). An increase in homozygosity also has implications for parasite resistance (see section 2.3 above). In humans however, assortative mating appears to be for complex, polygenic traits such as intelligence (Epstein & Guttman, 1992). There is little or no evidence of homogamy for single gene effects (Cavalli-Sforza 1971). Due to this aspect of human patterns, theorists expect only a small increase in homozygosity in the population (Jensen 1978).

Although homogamy need not cause any change in gene frequencies in the population, assortment for certain traits could lead to differences in fitness between genotypes. This would result in less successful genotypes being selected against and cause changes in population gene frequencies. Sloman and Sloman (1988) propose that the differential fitness that results from assortative mating has been a major force in human evolution. They build an interesting theory on two main assumptions. Firstly, they assume that all humans within a population respond to the same criteria of attractiveness - a *type* rather than *homotype* preference. Secondly, those characteristics that are perceived as attractive are indicative of fitness (facial beauty at least does seem to indicate parasite resistance - (Thornhill & Gangestad, 1993). Sloman & Sloman claim that small variations in genetic fitness can be greatly amplified by social processes in dominance hierarchies. Individuals who are successful in competition gain a corresponding increase in status that would greatly increase inclusive fitness; "small genetic differences are writ large in social hierarchies". Those at the top of a hierarchy attract the best quality mates (and thus assort on the common criteria of 'attractiveness'), leaving those below to mate assortatively with other partners, as they have no other choice. A final assumption that status is related to the survivability of offspring leads Sloman and Sloman to the conclusion that small genetic differences can lead to greater representation in the gene pool due to status effects, speeding the process of evolution.

To summarise, assortative mating for genetically transmitted characters appears to occur in human populations. There are many possible advantages and costs of homogamy, and theorists are as yet undecided on which apply to humans. The arguments concerning genetic similarity in human marriage are still ongoing (Russell & Wells, 1994,1995; Rushton 1995). Before coming to any conclusions about the importance of assortment in human marriage and to try to separate type from homotypic preferences, we must look at possible mechanisms that would allow detection of genetic similarity to occur.

6.3 Detection of genetic similarity: influence on mate choice.

Having considered the theoretical implications of assortative mating and optimal outbreeding in the previous section - the 'ultimate' causes in terms of inclusive fitness – sections 6.3 and 6.4 will deal with the proximate mechanisms by which human and non-human animals develop sexual preferences. For optimal outbreeding to occur, individuals must detect the likely genotypic similarity between themselves and possible partners, and 'assess' whether they are too closely or too distantly related. The subject of kin recognition has been studied in some depth, but mainly from the perspective of kin selection and the evolution of altruism rather than optimal mate choice (Hamilton 1964). However, as information gleaned from kin recognition processes may also be used in mate choice to avoid excessive inbreeding (Bateson 1983; Waldman 1987), some proposed mechanisms will be briefly discussed. Following this discussion, evidence for sexual imprinting on self-similar individuals in non-human animals will be considered. Particular attention will be paid to Bateson's work with Japanese quail. Section 6.4 will review evidence for a similar imprinting process occurring in humans.

6.3.1 Kin recognition & incest avoidance among animals.

The first and simplest form of kin recognition is based on spatial cues. Especially in relatively static populations, individuals in the immediate vicinity are likely to be closely related. Waldman (1987) notes that even in mobile species behaviour (be it altruistic or sexual) may be a function of distance from some specific location such as a burrow or nest. Dispersal at sexual maturity is common in many species, including primates (Greenwood 1980, Holmes & Sherman, 1983, Packer 1979). Although inbreeding avoidance may not be the primary cause of such dispersal in all cases, it is one consequence (Bateson 1983).

Association or familiarity is probably the most prevalent form of kin recognition in many mammal and bird species (Holmes & Sherman, 1983). Repeated social interaction between proximate individuals (e.g. nest mates) at early stages of development seems to foster altruistic acts whilst inhibiting sexual interaction (Holmes & Sherman, 1983; Waldman 1987). Familiar individuals are assumed to be family. In a captive population of stumptail macaques who were prevented from dispersing as they would in the wild, Murray & Smith, (1983), found that close mother/son association was sufficient to cause a behavioural aversion to incestuous pairing. Experimentally, the association hypothesis of kin recognition is testable by placing infants in surrogate families at appropriate stages of development, and observing kin directed behaviours to these non-related 'off-spring'. Errors of kin identification in such circumstances have been observed in many species including fish, mice, and goats (see Holmes and Sherman, 1983 for a review). Familiarity is thought to be one of the primary mechanisms of incest avoidance in humans (the Westermarck hypothesis; see below).

Kin identification without social interaction (i.e. the recognition of relatives that have never been met) requires a different mechanism. Two have been proposed; recognition alleles and phenotype matching. Recognition alleles were first proposed by Hamilton (1964), who hypothesised that such alleles would simultaneously cause the expression of a phenotypic trait in an individual and cause beneficial behaviours to other individuals with that trait. This 'green beard' type process (Dawkins 1976) is considered unlikely by theorists (including Hamilton himself) due to the likely intergenomic conflict that would arise from their existence. If one such allele exists, others may coexist in the same organism potentially causing opposing behaviours (Dawkins 1976; Holmes & Sherman, 1983; Waldman 1987).

A phenotype matching process of kin recognition may involve learning of an individual's own phenotype and/or the phenotypes of kin. Putatively, these learnt phenotypic characteristics form a template to which unfamiliar individuals can be compared when they are first encountered. Individuals who closely match the template are more likely to be identified as kin than those who do not (Holmes & Sherman, 1983; Waldman 1987). Evidently, such a mechanism relies on a consistent correlation between genotype and phenotype. Phenotype matching has interesting implications for the detection of relatedness, and therefore optimal mate choice. This concept will be discussed with relation to Bateson's experiments with quail later. An innate disposition to learn the characteristics of kin for the purpose of future mate choice implicates some kind of innate learning process during development. Positive assortative mating requires a mechanism by which individual animals recognise and show a preference for individuals that are phenotypically similar to themselves, (and by extension similar to their close relations). Ethological studies of sexual imprinting provide a likely candidate for such a mechanism; these will be discussed below.

6.3.2 Imprinting on familial characteristics in non-human animals.

Many theorists have proposed that imprinting at an early age influences later behaviours including mating preferences in many species (Immelman, 1975). Examples of cross-species mating preferences have been demonstrated to be the result of imprinting on a unsuitable model (D'Udine & Alleva, 1983). An imprinting-like mechanism may play a crucial role in both incest avoidance and the detection of familial characteristics that allow positive assortative mating with non-family members.

Immelman (1975) reiterates Lorenz's primary properties of imprinting. Firstly, imprinting takes place during a specific period of an animal's life known as a 'sensitive period'. The length of this period varies between species and the behaviour that is imprinted, but the important concept is that experiences can have more effect on an individual's later behaviour when they occur during one period of development than another (Bateson 1978). Secondly, imprinted behaviours cannot be forgotten. Finally, imprinting can be completed long before the behaviour is implemented. Sexual imprinting is a good example of this phenomenon; preferences imprinted in infancy manifest themselves at sexual maturity, long after the end of the 'sensitive period'.

The best evidence for sexual imprinting for familial characteristics comes from studies of the development of sexual preferences of birds, particularly Japanese quail (Bateson 1978, 1980, 1983,1988; Ten Cate & Bateson, 1989), and the Zebra finch (Vos 1994, 1995a, & 1995b).

David Vos' work with Zebra finch (Vos 1994, 1995a, 1995b) indicates that parents have an influence on partner choice, and that there may be sex differences in imprinting, both from the chicks perspective (male and female chicks imprint on different characteristics) and the parents (infants may preferentially imprint on one sex parent or the other).

Vos studied the preferences of male zebra finches by mating birds with different morphs (Vos 1994). He found that males showed a sexual preference for a female of the mother's morph than the fathers. Moreover, most of these male birds showed a preference for a *male* bird of their mother's morph over a female finch of their father's morph. A further study of zebra finch sex differences in imprinting exploited the fact that white zebra finches lack any sexual dimorphism except for a difference in bill colour - female beaks are orange and male beaks are red (Vos 1995). Vos manipulated the appearance (and therefore the apparent sex) of mating birds by painting their bills with either red or orange nail varnish just before their chicks opened their eyes. This gave him two groups of test birds; one group raised by 'natural' coloured parents (i.e. mother's bill painted orange, faather's red) and another group that had gender switched characteristics (father's beak orange, mother's red). Stimuli birds of both sexes with both beak colours were prepared for testing. Testing involved presenting pairs of stimuli birds to the test birds in separate side cages. Stimuli birds representing all possible combinations of sex and bill colour were presented in this way. The results showed that female finch preferred to associate with male rather than female birds (even in test trials in which opposite sex stimuli birds had the same colour beak), and beak colour had no effect on their choices. It should be noted that no one way screens were used in Vos' apparatus - stimulus birds could see the test birds. Hence successful sex recognition by female finch could have resulted from behavioural cues. The sex and the bill colour of stimulus birds had significant effects on the preferences of male finch. - they preferred birds with the same bill colour as their mother, and also showed an independent, but weaker preference for females. Vos studied the preferences of female finches further (Vos 1995). By using mixed morph pairs to rear infants he found that females also developed preferences for males with their mother's plumage. Bateson's seminal work on optimal outbreeding with Japanese quail is related to Vos' studies, but in addition demonstrates the concept of optimal inbreeding. A discussion of this work follows.

6.3.3 Imprinting and optimal outbreeding in Japanese quail.

Bateson's 1978 study was the first to demonstrate that quail develop preferences for novel individuals who are similar to birds that they are familiar with. Quail were raised in randomly selected groups of 10 chicks. Males were isolated from other chicks at 35 days of age, and were tested at approximately 60 days, when sexual maturity was reached. Preferences towards familiar and novel birds were tested; similarity was manipulated by using female quail with either white or brown plumage. Two tests were performed. The approach test allowed male quail to approach two female birds that were behind separate Perspex screens. When given a choice between a familiar bird (with whom the test male had been reared) with brown plumage and a novel brown female, males showed a strong preference for the novel bird. When choosing to approach either a novel brown or a novel white bird, males once again preferred the unfamiliar brown. The second test involved placing a male in a cage with two female birds, and recording which female was the target of the first ten attempts at copulation. Once again males preferred a novel to a familiar brown as a partner. However, the familiar brown was preferred to a novel white bird. Whilst these results could be interpreted as demonstrating a male preference for brown over white females, earlier literature indicates that this is unlikely (Gallagher 1976). Instead, Bateson postulated the preferences for novel over familiar brown plumage birds was caused by males selecting an optimally discrepant mate - the white birds were too dissimilar to the birds that the male was reared with to be considered a viable mate (figure 18). The use of a two way choice test is ambiguous - the response of the test bird will depend on how much the familiar and novel stimulus birds differ (i.e. how far along the x axis of figure 18 the novel bird is found). It can be seen from the figure that very high or low similarity will cause no preference for novel birds. The copulation test generated very similar results to the approach test, indicating that time spent near a opposite sex bird could be interpreted as a sexual preference. This inference enabled Bateson to use just approach tests in his later studies, which are easier to control than copulation tests.

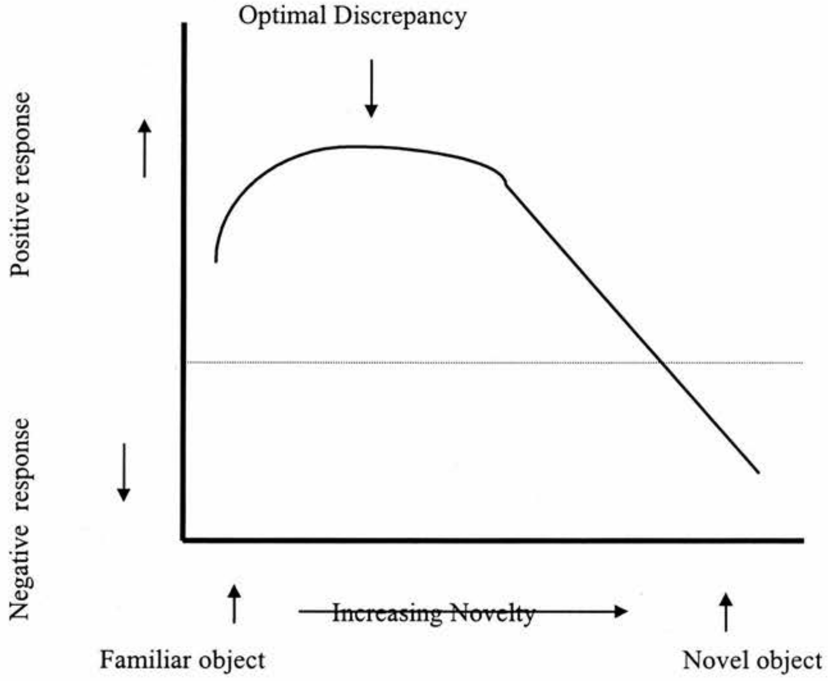


Figure 18. Bateson's proposed relationship between familiarity and sexual response. In the case of his 1978 study, the familiar object represents a brown morph female with whom the test male had been reared, the novel object is an unfamiliar white morph female, and the unfamiliar brown female represents an approximation of an 'optimally discrepant' mate. (From Bateson, 1978.)

Bateson (1980) continued to examine the preferences of Japanese quail when they were presented with a familiar bird with which they had been reared, and a novel quail unknown to the test bird. Test birds were reared in groups of ten for 30 days (the other members of the group were not necessarily siblings) before being socially isolated. The birds were again tested at around sixty days of age. Two stimulus birds were placed in the testing apparatus behind one way screens, so that they could not visually interact with the test bird. Bateson recorded the time the test quail spent in front of each of the stimulus birds as a measure of sexual preference. Overall, male quail showed a significant preference for novel birds rather than familiar birds. A weak, non-significant female preference for novel birds was also found. The male preference for novel over co-reared individuals is reminiscent of the sexual aversion humans appear to feel for co-socialized peers, as detailed above. Bateson's interest in how early experience leads to later preferences prompted him to study the composition of the rearing groups. He hypothesised that the number of opposite sex birds in the group may be important, and so analysed the data separately in groups dependent on the number of opposite sex conspecifics. He found that for male birds, a significant preference for novel birds was shown only by birds that had been raised with two females in their rearing group. Conversely, females raised with two males in their group spent significantly more time in front of the familiar stimulus bird. Females raised with three males spent significantly more time with novel birds, and all other group compositions lead to non-significant preferences towards unfamiliar. So a mixed pattern of results is found; although overall preferences are for novel birds (as Westermarck would predict for humans), these results appear to be dependent on both the sex of the test bird and the composition of the rearing group. Bateson concedes that the data are not ideal for analysis; the death of some birds during rearing lead to different group sizes that prevented sophisticated statistics. The analysis of varying group composition consisted of ten one sample t-tests that generated three significant results - the 50% probability of a type one error here should not be overlooked.

Nevertheless, the findings are theoretically interesting and led Bateson to develop a model of optimal preference development in which the number of opposite sex individuals in the rearing group plays a pivotal role. The characteristics of each co-reared individual provide a 'zone of familiarity', which serves as a model to which potential partners can be compared. Bateson assumes that the ideal mate is slightly different to familiar birds. This model is illustrated in the figure and accompanying legend below. The model could successfully account for the anomalous preference for familiar males that some female birds demonstrated.

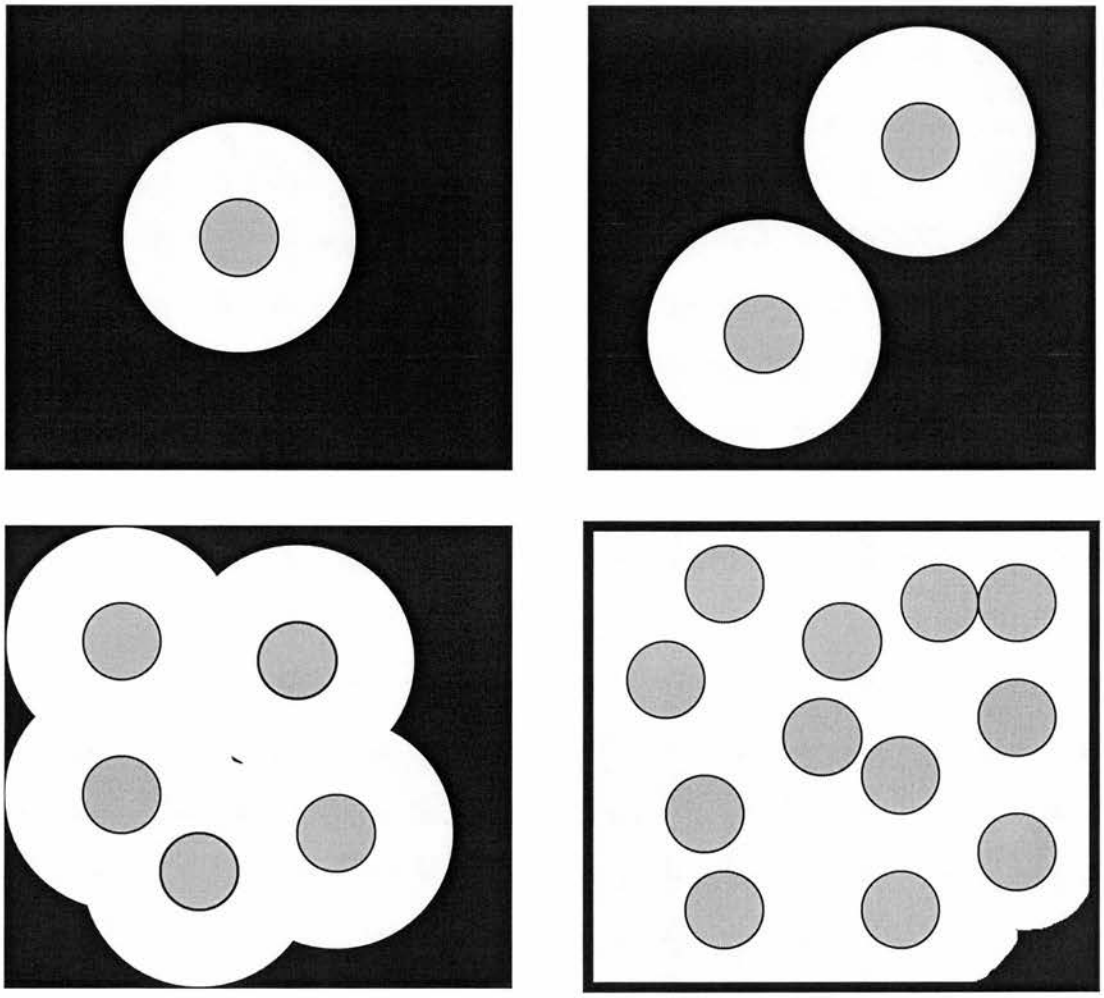


Figure 19: Bateson's 1980 model of optimal preference development. Small dark grey center circles indicate an zone of familiarity provided by one member of opposite sex in rearing group. Large lighter circles indicate preferred mate zone of optimal difference. The black area represents the zone of rejection - individuals falling within it are considered too different to the familiar to be suitable mates. The rectangle indicates limits of variation in external appearance in the population. Bateson's assumed rule is mate in light grey area if possible, but pick 'dark grey' rather than 'black' if 'light grey' is not available, i.e. to inbreed rather than hybridize. When number of co-reared opposite sex offspring is low (top left quadrant), chances that a novel individual will lie in the region of rejection is high, leading to a preference for familiar individuals. If number of co-reared birds is high (bottom right quadrant), the chances of novel birds falling in the black area (rejection) is low, but the chances of birds being too familiar (dark grey circles) is high. The condition leading to preferences for novel birds is somewhere between the two extremes (e.g top right, bottom left). (From Bateson, 1980).

Bateson's 1980 paper successfully showed that quail preferred to associate with unfamiliar birds, and he interpreted this as a sexual preference. However, whilst demonstrating that quail prefer not to inbreed where possible, the study fails to demonstrate that the quail are seeking any similarity between themselves and a novel partner. Bateson's 1982 study attempted to test the optimal outbreeding hypothesis that the preferred level of relatedness that quail choose in a partner would lie between close relatives and unrelated birds. As the pedigree of the quail in Bateson's laboratory population was known, birds with various levels of relatedness could be selected for mate choice experimentation. Test birds were raised in their true families, not arbitrary groups as in the 1980 study. At maturity, the test quail were placed in apparatus that allowed test bird to observe five stimuli birds through one way screens, without interaction. The five birds (familiar sibling, novel sibling, 1st cousin, 3rd cousin and an unrelated quail) were chosen by their relatedness and familiarity with the test bird. The use of five stimulus birds overcame the ambiguity of two stimuli choice tests mentioned above. The sample of 22 male and 13 female quail spent significantly more time in front of 1st cousins than unrelated birds and novel or familiar siblings. The difference in time spent in front of 1st and 3rd cousins is unreported, but presumably non-significant. Although there were no significant sex differences, the individual data were highly variable - only 12 of the 35 five birds spent most time near a novel first cousin, a further nine preferred the 3rd cousin. As the coefficient of relatedness between two birds is the probability that those individuals will share any particular gene, closer relations will *on average* resemble each other more than distant relations. However, in any experimental trial the unrelated stimulus bird may resemble the test bird more than the 1st cousin stimulus bird, which Bateson proposes to be the cause of the variability in the data.

The 1982 paper represented a significant theoretical step in studies of optimal outbreeding - it was the first empirical evidence that birds avoid both excessive inbreeding and outbreeding. Bateson (1988) proposed that the proximate mechanisms of this behaviour are imprinting and habituation; quail imprint on family characteristics generating preferences for self-similar birds, but then habituate to individual family members, rendering them less arousing than non-familiar birds with similar characteristics. Batesons's 1988 study went further, and examined fecundity of pairings between birds with varying levels of relatedness. As before, quail were raised with true siblings for 35 days. At 40 days, the birds were paired in a separate pen with an opposite sex bird that was either a familiar sibling, a novel first cousin or a novel unrelated bird. The age at which these pairs laid their first egg and their first fertile egg were noted. The number of eggs laid in three periods after hatching was also recorded.

The results showed that 1st cousin partnerships produced fertile eggs significantly before sibling and unrelated pairs. Initially, cousin partnerships also laid significantly more eggs than sibling or unrelated pairs, but these differences disappeared with time. Bateson suggests that the difference in laying of fertile eggs is not due to genetic incompatibility (witnessed by the eventual similarity in laying rates), but preferences - quail mate preferentially with 1st cousins in Bateson's laboratory population, supporting the findings of his previous studies, and concluding what are so far the most convincing experimental studies of optimal outbreeding.

Bateson's and Vos' studies are interesting in that they point out the potential complexities present in preference development. Sexual imprinting in any species is likely to involve complex interactions between the exposure to parents, siblings and the sex of the individual learning the behaviour. Natural conditions are never employed in experimental studies of imprinting, which may have a profound influence on the behaviours demonstrated by laboratory animals (Lickliter & Gottlieb, 1985). Nevertheless, for the purposes of the current study evidence that early experience can have predictable effects on future sexual preferences in many animals seems sufficient to stimulate a search for similar effects in humans. The next section attempts to achieve this.

6.4 The Westermarck hypothesis, imprinting and incest avoidance in humans.

To assess the likelihood of optimal outbreeding in humans, the mechanisms that allow an optimally similar partner to be chosen must be elucidated. To avoid inbreeding, an aversion to sexual interaction between consanguineous individuals must occur, and yet for positive assortative mating to occur due to preferences (rather than, say, propinquity) individuals must recognise similarity in prospective mates. The evidence for an innate disposition to 'imprint' incest avoidance in humans is discussed below (Westermarck's hypothesis), followed by a review of studies of sexual imprinting on familial characteristics.

The most likely mechanism to avoid sibling and parent sibling incest can be found in the Westermarck hypothesis. From his studies in Morocco in the late 19th century, the anthropologist argued that children have an innate tendency to develop a sexual aversion to individuals with whom they live closely in infancy and early childhood, i.e. siblings and parents. (Westermarck, 1894). Although his theory was greeted favourably at the time, by the twenties the leading scholars of the day (including Freud, Durkheim & Malinowski) had rejected Westermarck's hypothesis (Wolf, 1993), and proposed that a culturally imposed taboo was the major prevention of incest, and one of the defining characteristics of humanity. The "social science orthodoxy" that replaced Westermarck's theory of incest avoidance was built on erroneous assumptions; critically that incestuous mating is the 'natural' choice throughout the animal kingdom (and would be in humans without the intervention of culture), and that incest taboos are universal in humans (van den Berghe, 1983). Incestuous matings are clearly not the norm in non-human animals, and are avoided by various mechanisms such as dispersal or possibly a Westermarckian type aversion (Murray & Smith, 1983, see also van den Berghe, 1983). Excellent discussions of the strengths and weaknesses of Westermarckian and cultural theories of incest avoidance can be found in Wolf, 1993, and McCabe, 1983. Since the fifties, a series of ethnographic studies have formed 'natural experiments' that seem to test, and support, Westermarck's opinion. These studies share the common feature that infants and children are placed in intimate contact with other infants in a sibling like relationship, and yet they are encouraged (or in some cases not prevented) to interact sexually later in life with these co-socialized individuals. Details of the three of these studies are given below.

6.4.1 Kibbutz studies of Westermarck's hypothesis.

The studies most cited in support of Westermarck have taken place in kibbutzim in Israel (e.g. Shepher, 1971; Spiro, 1965). The methods of child rearing in Kibbutzim provide an ideal test of the Westermarck hypothesis. Firstly, children are socialized from infancy in a mixed sex peer group (a *kitah* or *kevutza*), determined by age. They live separately from their parents in a kitah (although parents spend time with their children in the afternoons), in which group solidarity and cooperation between *sabras* (comrades) is encouraged. Communal eating, sleeping, bathing and play arrangements mean that sabras spend most of their time together. The trained caregivers do not interfere with sexual play, and at no point during infancy or early childhood is sexual interaction between sabras formally prevented. Thus, a situation is generated in which the relationship between sabras has much in common with siblings in other cultures (a lengthy period of intimate contact with other individuals during childhood) but also crucial differences (most importantly, the lack of a culturally imposed incest taboo). Westermarck's hypothesis predicts that despite this lack of an incest taboo, the innate predisposition against sexual interaction with co-socialized peers will prevent sabras interacting sexually. Overwhelming evidence appears to support Westermarck, as detailed below.

The findings of Shepher's detailed study of one kibbutz (Shepher, 1971) are easily summarized; there were no cases of heterosexual activity or marriage between members of the same kitah. Shepher reports that the lack of sexual attraction between sabras and the subsequent voluntary avoidance of sexual interaction is actually regretted by those involved. Shepher also summarizes impressive census data in support of Westermarck: out of 2769 marriages involving individuals raised in kibbutzim "there is not a single case of true intra-peer group marriage to be found" (p.297). Although 14 cases of marriage between peers were reported, follow up studies revealed that none of these couples had been in the same kevutza for more than two years before the children were aged six. This implies that there may be a 'critical' or 'sensitive' period in which the sexual aversion develops - an important point that I will return to later.

Shepher's studies are not alone in finding a lack of sexual interaction between kibbutz peers. Rabin, 1965, Bettelheim, 1969, and Talmon, 1964, all report similar conclusions from their kibbutz studies. Spiro, who performed a lengthy study of kibbutz life in the early fifties, concedes that "students who have always been members of the same *kevutza* have never been known to engage in sexual behavior with each other" (Spiro, 1965, pg.334), despite his disbelief in Westermarck's thinking. Critics of Westermarck point out that the compulsory two year period of military service in Israel coincides with the time that children would be expected to find future partners (ref from van de Berghe commentary), and that this alone may prevent kibbutz children from marrying peers. However, this argument fails to address the lack of adolescent sexual interaction witnessed in kibbutz life.

6.4.2 *Sim-pua* marriage in China, and other evidence for Westermarck's hypothesis.

The second commonly reported body of data in support of the Westermarckian thesis is reported by the anthropologist Arthur Wolf in his studies of *sim-pua* (minor) marriages in Taiwan (Wolf 1966, 1993; Wolf & Huang, 1980). Until the mid-1940s women in much of China and Taiwan gave away their infant daughters, and instead raised other women's daughters as future wives for their sons. The future husband and wife are raised together, sharing sleeping mats in the same fashion as siblings. Westermarck would predict that at sexual maturity, the husband and wife should have developed an aversion to consummating the marriage due to their intimate childhood association.

Supporters of Westermarck consider that the Chinese data neatly fill important gaps in the Israeli literature. Firstly, kibbutz children are never forced to marry peers - the aversion they feel to each other may be depthless, and easily overcome once sex has been made a part of the previously platonic relationship. In the Chinese case the childhood associates are compelled to marry by extremely strong social pressures. An excellent natural control group is available to compare the relative success of these forced unions to marriages without co-socialization of partners; the alternative to sim-pua marriage was to marry a mature son to a post-pubescent girl from another family in 'major marriage'. Such couples had often not met until the marriage ceremony, thus precluding any aversion to sexual activity. Wolf reports in his original fieldwork (1966) that childhood associates are extremely reluctant to consummate their legal bond - the marriage ceremony is itself referred to as "pushing them together". Husbands in sim-pua marriages were three times as likely to visit "dark rooms" (brothels) as men married in the major fashion, an observation that Wolf interpreted as a lack of satisfaction with the sexual aspect of their marriages. Later census studies of more than 14,000 marriages indicate that 2.5 sim-pua marriages end in divorce for every major marriage that does so, and the fertility of major marriages is 25% higher than minor marriages (Wolf 1993).

Wolf's census data also answer many of the objections to his data (Wolf 1993). The differences in fertility between marriage types are unlikely to be caused by 'adoption trauma' or maltreatment of the adoptee. Girls raised for a sim-pua marriage that did not occur (often due to the death of the prospective husband), but instead married other men, had the same fertility and divorce rates in their marriages as women in major marriages. In addition, Wolf ruled out links between socio-economic status and the low fertility of minor marriages (poorer families tended to prefer minor marriages), by considering the taxable value of the husband's estate in his analysis.

Interestingly, Wolf also claims to have discovered a 'critical period' in which the aversion between sim-pua marriages develops. He notes that the age at which the bride moves into the husbands home is crucial; "Brides adopted before age three display sharply reduced fertility and a markedly higher probability of divorce; those adopted after age three do not." (Wolf, 1993). When he reanalyzed the data considering only minor marriages in which the brides moved into the husbands residence before three years of age, he found that the relative minor/major marriage divorce rate increased to 3:1, and that major marriages were 45% more fertile.

The final empirical support of the Westermarck hypothesis is McCabe's study of father's brother's daughter (FBD) marriages in Lebanon (McCabe 1983). She noted many similarities between the relationships of opposite sex siblings and those of first cousins; both were characterised by constant intimate interactions such as eating, playing and sleeping together at a young age. Later in childhood, McCabe describes the informality and candour between opposite sex first cousins using similar terms as Spiro, Shepher and others used when describing Kibbutz relationships. Hence, McCabe hypothesized that should these cousins be partners in FDB marriage (which is culturally prefered), a Westermarckian impediment may prevent successful relationships developing. In a similar (though smaller scale) analysis to that performed by Wolf in Taiwan and China, McCabe studied fertility in such FDB pairings and other pairings in which co-socialization had not taken place. She found that FDB marriages produced 23% fewer children and a fourfold increase in divorce rates when compared to other pairings in her sample, in support of the Westermarck hypothesis.

The evidence above, coupled with other non-empirical studies of Westermarck's thesis e.g. Burling, 1985, seems to indicate that childhood association causes later sexual disinterest between peers. Other forms of incest (e.g. mother-son, father-daughter) have not been discussed here, but arguments extending Westermarckian theory to cover these cases can be found in Wolf (1993), McCabe, (1983) & van den Berghe (1983). Interestingly, theorists seem to propose a 'critical' or 'sensitive' period of life during which co-socialization must occur to develop these aversive sexual feelings towards certain individuals later in life. Most scholars have been quick to point out that such a critical period is analogous to imprinting processes in non-human animals (Shepher, 1971). So an innate aversion to incestuous sexual activity seems likely. However, optimal outbreeding theory also requires that individuals recognise and are attracted to self similar phenotypic characteristics in potential non-incestuous partners. The next section will cover human evidence for possible imprinting on family (and therefore by extension self similar) characteristics.

6.4.3 Human imprinting on familial characteristics.

Evidence for an imprinting type process of learning leading to preferences for familial characteristics in human partners is limited. Three studies of facial similarity between partners have already been discussed above (Griffiths & Kunz, 1973; Hinsz 1989; Zajonc et al, 1987). The findings of these studies could be attributed to the development of preferences for family-like (i.e. self-similar) facial characteristics in early life, although this is by no means the only possibility. Indeed an imprinting based explanation fails to explain the idiosyncrasies of the data generated, such as Zajonc's findings of a reported increase in partner similarity with age. Nevertheless, two other studies indicate that parental characteristics, and specifically opposite sex parental characteristics, could influence later choice of partner.

The first of these studies is an analysis of the tenth Italian census data (Zei, Astofli & Jayakar, 1981), which investigates the relationship between father's age and husband's age for a sample of over 350,000 women. Small but consistently positive correlations between these variables indicate that the daughters of older men subsequently tend to choose older husbands. Separate analyses were performed for two separate time periods (a fifteen-year pre-war and fifteen-year post-war period) and for different locations around Italy. The age of the women at marriage and their educational level was also taken into statistical consideration. Correlations in all analyses were positive, but were found to be stronger in rural and less economically developed areas. The overall positive correlation between father's and husband's age is consistent with the possibility of an imprinting type process; daughters may imprint on the characteristics of their father (including their age) as children and later find these characteristics preferable in their own partner. James (1983) criticized this study on the grounds that the age of the father at the birth of his first child is correlated with social class. Hence he proposes that Zei et al's correlation is likely to be caused by social class, not imprinting. Zei et al (1983) reply that their control for educational level must to some extent also control for social class, and reject James' criticism. They point to the stability of their findings across different time periods, locations and living styles across Italy as convincing evidence that father's age does indeed correlate with husband's age irrespective of social class.

A second study provides further evidence for a correlation between father's age and partner's age. Wilson & Barrett (1987) obtained responses to a questionnaire from 314 young British women, giving data on the ages and the eye colour of the women, their partners and their parents. The age data showed a correlation of borderline significance between father's age at the time of the daughter's birth and the daughter's partners age; a result consistent with Zei et al's findings. The eye colour data collected showed a weak non-significant trend for girls to choose boyfriends whose eye colour resembles their father. One hundred and ninety seven respondents had parents with differing eye colour. Ninety-nine of the girl's partners had eye colour similar to the father, compared to 87 partners who resembled the mother; 11 subjects were discarded. Another analysis was performed on a subset of data based on subjects who reported their partners and parents as having blue or brown eyes. Wilson & Barrett justified this seemingly arbitrary filtering of the data on the grounds that these were the most distinguishable categories. 'Intermediate' categories such as green were discarded, presumably to reduce noise in the data. The results of this analysis indicate that the girl's partners were significantly more likely to have the same eye colour as their father than the other colour (from the blue/brown dichotomy). The respondents also tended to choose partners that had their mother's eye colour, although this trend was non-significant.

Although both the Italian census data and Wilson & Barrett's work are encouraging to those looking for evidence of sexual imprinting in humans, it should be noted that the significant correlations and tests found in both studies have very small effect sizes, and account for very little of the variance in the data (the correlations between father's and husband's age in Zei et al's study are between 0.05 and 0.08; in Wilson's study the same correlation is 0.11) . The difficulties in detecting the influence of any of the multitude of factors that contribute to mate choice is clear - very large samples are needed to achieve statistical credibility when any one factor is considered.

6.5 EXPERIMENT 6

6.5.1 INTRODUCTION

The study of similarity of real life couples has obvious validity in the study of assortative mating for facial appearance, but also has drawbacks; studying partnerships is not the same as studying preferences (Burley, 1983) and studies can have confounding factors such as a common source of photographs for partners.

A computer graphic study is well placed to investigate assortative preferences for facial characteristics. Experiment 1 employs techniques that change the apparent sex of an individual's face, while maintaining their own characteristics (Rowland & Perrett, 1995). This creates an image of a hypothetical opposite sex 'sibling'. As there can be no Westermarckian impediment to an individual finding a synthesized facial image attractive, these sibling images can be used as test stimuli in studies of similarity and facial attraction.

6.5.2 METHOD

STIMULI.

52 female and 23 male participants (students at St Andrews University, mean age, 21) were photographed and the images digitised. The positions of 174 feature points were marked on each image to define the shape of the eyes, mouth etc. Component images were then blended to form average or prototype images (Benson & Perrett, 1992; Perrett et al., 1994; Rowland & Perrett, 1995; Figure 20).

Figure 20: Female and male averages used for transformations in experiment 7

An opposite sex image that retained shape information from the individual source faces was generated for each of the 52 female participants photographed. The vector difference between the feature points of an individual (e.g. female subject) and the same sex (female) prototype specifies the shape information unique to the individual. This identity information can be added to the shape of the opposite sex (male) prototype to create a synthetic male with a 'similar' face type to the subject (Figure 21, left vs. centre, Rowland and Perrett, 1995).

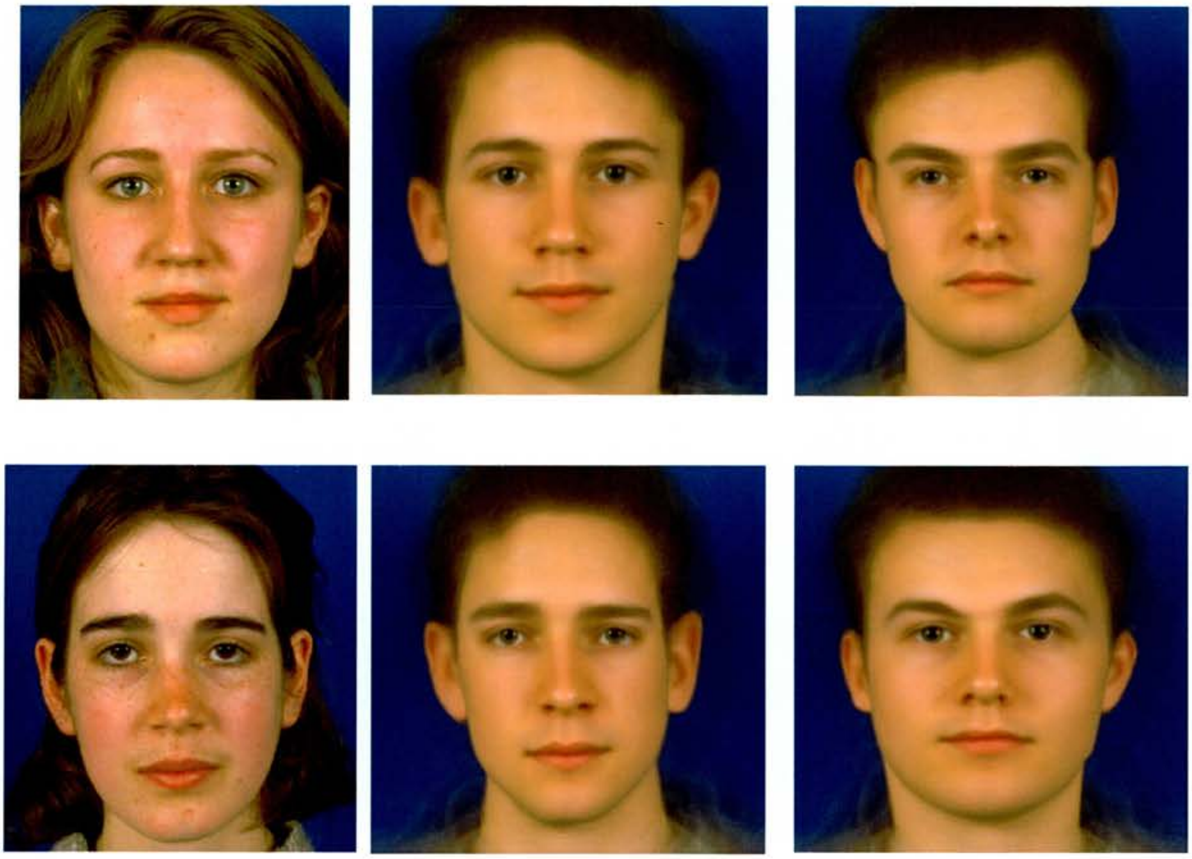


Figure 21: sex transforms of two individuals: Examples of original faces and synthetic stimuli. Original female (left), 'similar' male face stimuli (centre, experiments 1 & 2) and 'opposite' male face stimuli (right, experiment 2)

ATTRACTIVENESS JUDGEMENTS.

36 of the original 52 female subjects photographed rated the attractiveness of 52 transformed 'male' faces on a 7 point Likert scale (1 = 'very unattractive', 7 = 'very attractive'). The order of presentation was randomised.

SIMILARITY ASSESSMENT.

6 different subjects rated the similarity (on a 6 point Likert scale) of each of the synthetic male faces to each of the original photographs of the 36 female subjects making attractiveness judgments. Each subject rated the similarity of the complete set of 52 male faces to one original face before proceeding to repeat the task for a different original face (randomly chosen from the 36). Ratings were self-paced in sessions spaced over several days.

6.5.3 RESULTS

None of the subjects spontaneously recognised the stimuli as being derived from the face of themselves or their peer group. This might mean that the attempt to construct stimuli similar to subjects was unsuccessful. Analysis of similarity ratings showed this was not true. For each original female face the 52 male faces were ranked by the average similarity rating across the 6 raters. For all but 4 of the 36 original faces, the male face constructed as an opposite sex version (Figure 21, centre images) ranked highest in similarity from the set of 52. Two synthetic male faces ranked fifth and 2 ranked second most similar to the original face from which they were derived.

Figure 22a plots the relationship between similarity of stimuli to the subjects and the subjects' judgments of stimulus attractiveness. Attractiveness and similarity ratings were Z transformed to allow comparisons among subjects. The 36 subjects' ratings of the male face most similar to themselves were averaged to give the extreme right data point. The extreme left point gives the average of the 36 subjects' ratings of the face least similar to themselves. Points in-between reflect the averaged attractiveness ratings of faces with intermediate rated levels of similarity. The point on the extreme right appears an outlier, with a much higher similarity level than the average similarity of the second most similar face. This is not an artifact but an indication of the success of the stimulus construction. The sex transformed stimulus was on average much more similar to the subject from which the face shape information was derived than any of the stimuli constructed to look like the 51 other female students.

The graph shows a clear relationship between similarity of faces to subjects and the subjects' ratings of attractiveness of the faces. Linear regression accounts for 58% of the variance in the data. Regression with a second order polynomial equation provides a better fit to the data and accounts for 66% of the variance. This indicates that although attractiveness ratings increase with similarity, this relationship asymptotes or declines when faces become very similar to the subject. This finding is consistent with the optimal outbreeding hypothesis, which postulates that intermediate levels of similarity should be most attractive.

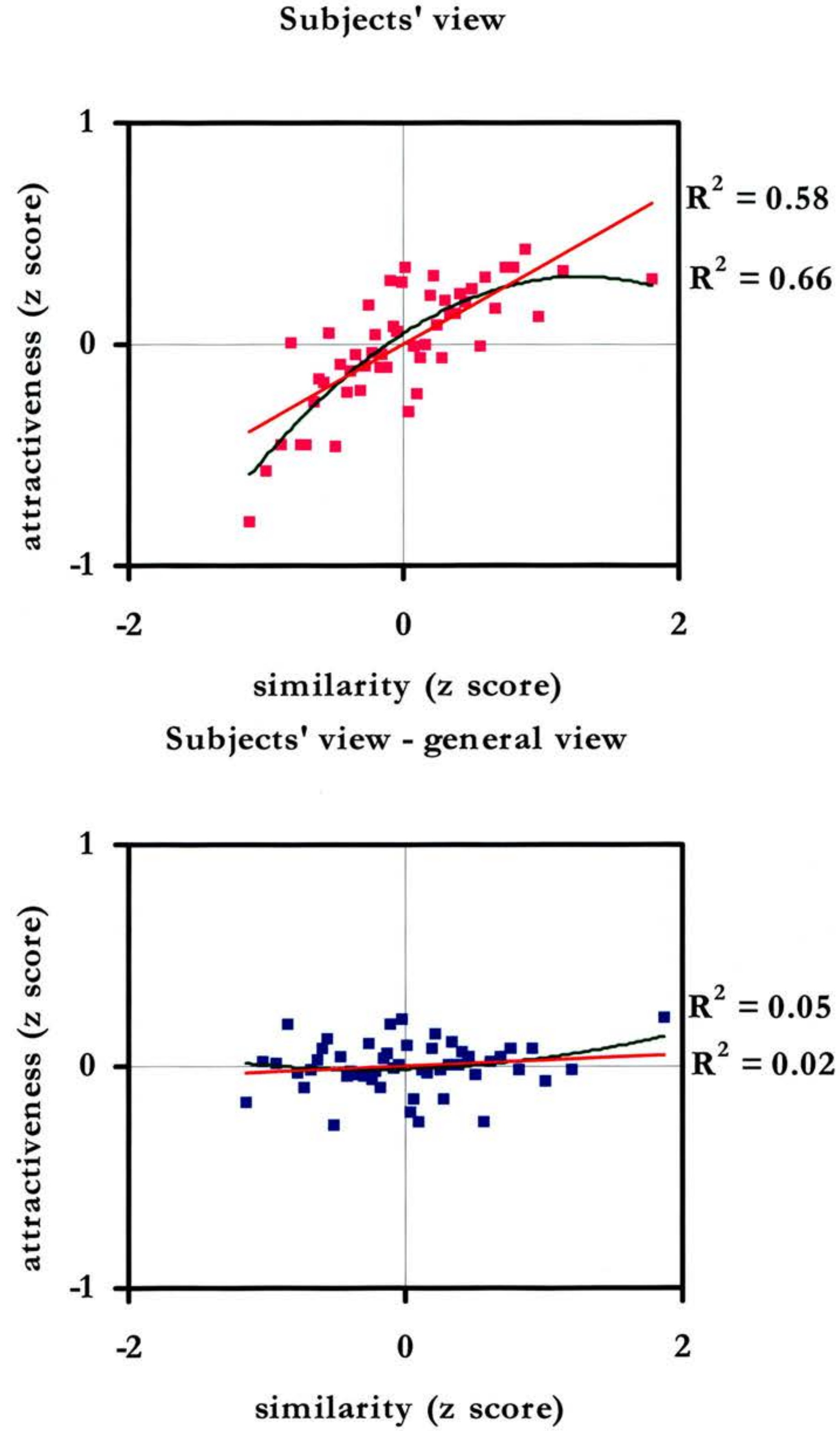


Figure 22: Facial similarity and judgements of attractiveness. A. (top panel) Average of attractiveness scores assigned by 36 subjects to 52 face stimuli ranked according to their similarity to each subject (for Method see text). Trend lines calculated by regression analyses using both linear (dashed line, accounting for 58 % of the variance in the data) and second order polynomial models (solid line, accounting for 66% of variance). B. (bottom panel) Subjects' 'unique view' of attractiveness, displaying the average difference between an individual's attractiveness rating of a face and the average ratings of all other subjects of that face. Linear and polynomial regression account for 2% and 5% of the variance respectively.

Figure 22b plots the attractiveness ratings of subjects in isolation from the views of others. A critical question is whether subjects rate self-similar faces differently to other members of the population. This can be assessed by taking the difference between each subject's ratings of a face and the average rating of everyone else for the same face. This value (subject's view - others' view) gives the subject's 'unique view' and is presented in Figure 22b. The graph displays no consistent relationship between the similarity of stimuli to subjects and their unique view of attractiveness.

The most exacting test of assortative preferences concerns the subject's ratings of the face most similar to themselves. This is analogous to testing whether the extreme right point in Figure 22b has a rating that is higher than zero. For each of the 36 subjects, two ratings were compared. The first of these was the rating given by a subject to the image rated as most similar to that individual (the 'self-similar' rating). In 32 of the 36 cases, this was the opposite-sex image generated from the subject's own face. The second was the median of all the other subjects ratings of that face (the 'others' rating). A Wilcoxon signed ranks test showed a non-significant trend for self-similar ratings to be higher than others ratings of that image ($Z=-1.692$, $p=0.091$, $n=36$).

6.5.4 DISCUSSION

Figure 22a indicates that (a) subjects are attracted to others with similar faces and (b) that preference peaks or asymptotes at a moderate level of similarity. Such evidence might be taken as support for the notion of optimal outbreeding and assortative mating. Figure 22b appears to contradict this.

It is not obvious how to reconcile Figure 22b with Figure 22a, in that it is not intuitive as to why subjects should be attracted to similar faces but no more so than other members of the population. The results can be explained in terms of a preference for average face shapes. Consider faces that are very different from average; highly unusual faces are likely to receive low ratings of attractiveness but, by definition, such unusual faces are unlike most people. Subjects are not attracted to faces that look very different from them but neither is anyone else. Now consider faces with an average shape that are viewed as attractive (Langlois & Rogman, 1990). A given subject's face will be more similar to an average face shape than to a face shape randomly chosen from the population. This means that average face shapes will have moderate levels of similarity to most subjects and will be rated relatively highly for attractiveness.

From these arguments it can be seen that optimal outbreeding, positive assortative mating and the averageness hypothesis all converge to make the same prediction: that like will prefer like (i.e. subjects will be attracted to similar looking others). While Burley (1983) notes that an assortative mating pattern need not be caused by assortative preferences, our data indicate that the corollary is also true: a *preference for similarity (i.e. averageness)* need not translate into a *pattern of similarity* in partners' faces. Consider a population of individuals of varying attractiveness (averageness) in which each individual competes to get the most attractive partner. The result is partnerships of equivalent levels of attractiveness (Berscheid et al., 1971). Whilst highly attractive individuals should end up with similar (average) looking partners, there is no reason for unattractive couples to look alike: each partner may have a face shape that differs from average in unique ways. Thus the averageness hypothesis predicts a preference for similar partners but game theory predicts that this will not translate into a pattern of physical similarity in partners (positive assortative mating).

Figure 22b provides a measure of attractiveness relative to the opinion of others. Subjects may find faces that are slightly similar to themselves (i.e. average faces) attractive and faces very different from themselves (i.e. faces far from average) unattractive, but other members of the population are likely to have similar views about these same faces. From this analysis subjects will not have extra motivation (above other members of the population) to seek out partners with a similar face.

In summary, Experiment 6 found evidence for a positive relationship between facial similarity and judgements of attractiveness across a range of face stimuli (Figure 22a). This preference could simply reflect the relationship between averageness and attractiveness. There was, however, a trend for an assortative preference, i.e. subjects rated faces with the shape most similar to their own slightly higher than other subjects.

6.6 EXPERIMENT 7

6.6.1 INTRODUCTION

A more direct test of positive assortative preferences is to compare a subject's ratings of an opposite sex face that is constructed to have the same shape as that subject and an average face of the opposite sex. Experiment 7 was designed to allow this direct comparison. In addition Experiment 7 addresses the concept of optimal similarity and investigates the possibility of negative assortative preferences for facial appearance. Negative assortative mating can be predicted from theories of parasite driven sexual selection since preference for partners with a dissimilar genotype should increase heterozygosity in offspring and thereby improve immunity (Thornhill & Gangestad, 1993). Human preferences for odour suggest negative assortative mating (Wedekind et al., 1995).

Novel interactive computer techniques allow warping between two different faces in real time giving subjects opportunity to select an optimum blend along a smooth continuum (Perrett et al., in submission). By constructing a continuum from an individual's sex transformed face through an average face to a face with dissimilar characteristics, the relative influence of similarity, averageness and dissimilarity can be assessed. Thus, three competing hypotheses based on evolutionary concepts of fitness can be tested. (1) Individuals will be attracted to some optimal degree of self-similarity in a face to realize the possible fitness benefits of positive assortative mating - the 'like prefers like' hypothesis. (2) Individuals will be attracted to facial shapes dissimilar to their own to maximise heterozygosity - the 'opposites attract' hypothesis. (3) Individuals will prefer average faces to self-similar or dissimilar face shapes as selection acts against extreme, non-average genotypes - the averageness hypothesis.

6.6.2 METHOD

SUBJECTS & PROCEDURE.

40 female and 21 male (mean age, 21) participants were photographed, and the images digitized as in experiment 1. The identity information used to create the similar face type (figure 21, centre) can be subtracted from the male prototype to create a synthetic male with an 'opposite' face type (Figure 21, right). If the original female had a small nose and thick lips, the similar male face would have both characteristics but the opposite male face would have a large nose and thin lips.

These two new face shapes (the similar and opposite face types) formed the end points of an interactive continuum, in which participants could manipulate the face shape displayed by moving a mouse controlled pointer left or right. Colour information from the appropriate sex prototype was rendered into the face shape in real time. Moving the computer mouse to the left or the right of the image showed the similar or the opposite face. Between these two points the image displayed a face shape in proportion to the position of the pointer. The centre of the range displayed a prototype image (50% of the similar and 50% of the opposite face shape - mathematically equal to the average). The interactive software ran on a Silicon Graphics Indigo^2 Maximum Impact workstation, in 24 bit colour. Two examples of the end-points of face continua are shown in Figure 21.

Seventeen females and six males took part in testing, all of whom had been photographed. Subjects were instructed to move the mouse left and right to view the continuum and press the space bar when the image on the screen was, in their opinion, most attractive. The software recorded responses in terms of the proportion of the transform selected by the subjects. After a short training period using continua of the same sex as the subject, the experimental stimuli proper were presented.

Male subjects performed 63 trials in total. The stimuli (female faces generated from the 21 male photographs) were grouped into three blocks of 21 trials; within each block the order of stimulus presentation was randomized. Similar and the opposite face shapes had an equal probability of appearing at the right or left end of the continua. Female participants performed 90 trials in total, in three blocks of 30 trials. Ten of the original 40 continua generated were not included in testing to reduce the duration of the testing. Order and left/right presentation was randomized as for male subjects.

Subjects were naïve as to the source of the stimuli; they were at no point informed that their own faces were used to generate the images.

6.6.3 RESULTS

The face shape selected by subjects in the continuum generated from their own facial characteristics is of primary interest. Subjects could select a face shape anywhere between a 100% similar to 100% opposite in shape. The mean level of similarity selected from this range was compared to the null hypothesis that the face shape selected would, on average, be neither similar nor dissimilar to the subject's own face (i.e. that the optimal face would be the prototype and 0% similar).

A one sample t-test against a hypothesized mean of 0% similarity preferred demonstrated that subjects did not generate responses that were significantly different from the prototype when presented with a continuum generated from their own face (mean preference = 2% similar, $t(22)=-0.44$, $p=0.67$). Separate analyses for each sex also yielded non-significant results (females, mean = 4.4% similar, $t(16)=-0.90$, $p=0.38$; males, mean = 4.4% opposite, $t(5)=0.35$, $p=0.74$).

6.6.4 DISCUSSION

Experiment 7 found no preference for any significant level of self-similarity in opposite sex partners. The design of this study placed it in a position to investigate the relative importance of two processes in preferences for facial shape: preference for average characteristics and preference for self-similar characteristics. Research has shown that average faces are in general more attractive than the individual faces from which they are synthesised, and this effect remains when skin texture and blurring are controlled in both original and average face shapes (Langlois & Roggman, 1990; Benson & Perrett, 1992). It appears hard to improve on faces with average characteristics. It may be that the fitness benefits associated with averageness (e.g. immunocompetence and fertility) far outweigh the hypothesised fitness benefits that accompany an optimal amount of self-similarity in a partner.

Our current experiments failed to reveal strong support for assortative preferences. One qualification of our studies is that they focused on female judgments of male attractiveness. Studies of sexual imprinting suggest stronger effects of early rearing on male subjects (e.g. Vos 1994, 1995a; Kendrick et al submitted). Previous studies of spousal facial similarity have reported sporadic positive results; one possibility for these findings is that subjects have matched partners on attractiveness rather than physical resemblance.

Perhaps a better test of assortative mating theory would be to allow subjects to choose between faces that differ from average by equal amounts but vary in similarity to the subject. If similarity between partners does correspond to a fitness advantage, and preferences for self-similar partners have evolved, subjects should prefer faces similar to themselves from a selection of equivalent non-average faces.

7 Computer Graphic Techniques

7.1 Averaging and blending: composite construction

Most of the graphic techniques in this thesis rely on the construction of facial averages or prototypes. Such averages represent the consistent attributes that a set of faces possesses. The use of composites stretches back to the work of Francis Galton (1878), who attempted to extract the facial characteristics of various groups of individuals such as criminals, army officers and people suffering various afflictions. Galton's technique aligned photographic negatives on top of each other before developing by manipulating their size and orientation so that the eye positions on each image were aligned. The resultant composite images represented the average properties of the constituent faces, but aligning on only two points meant that the final image was indistinct. Modern computer graphic techniques have allowed vast improvements over Galton's original technique. Langlois & Roggman (1990) adapted Galton's methodology using digitised images that had been stretched to align on both eye points and the centre of the mouth. This technique has also been criticised, as stretching the faces to align on the mouth may introduce artefacts (Alley & Hilldebrandt, 1991; Benson & Perrett, 1991). Aligning faces on three points leaves other features unaligned, and hence subject to blurring when the faces are blended digitally. This can make comparisons between averaged and non-averaged faces problematic.

The techniques described below improve over Langlois & Roggman's techniques in several ways: more points are used to mark feature points, allowing better alignment of features and less blurring. In addition, the use of many points allow the shapes of faces to be compared with some accuracy, facilitating caricaturing and transforms (such as the masculinity/femininity transforms used in expts 1-5, and the sex transfroms employed in expts. 6-7). All the software used to perform image manipulations runs on Silicon Graphics Maximum Impact Workstations; the techniques and images used are described below.

7.1.1 Delineation

To allow a single composite image to be generated from a set of faces, the shape information of each face in the set must be defined. Each face image is digitised, and then delineated using software running on a Silicon Graphics Maximum Impact workstation with 24 bit colour. To do this, 174 feature points are manually placed on the face, defining the shape of the eyes, lips, nose etc. (Benson & Perrett,1991). These feature points are shown in Figure 23.

7.1.2 Calculating average shapes

From the delineation data of all the faces in the set, the average position of each feature point on the face can be calculated. This involves calculating the mean position of each of the delineation points across the faces in the set that is being averaged: The mean XY position of each of the 174 feature points defines the average face shape of the set.

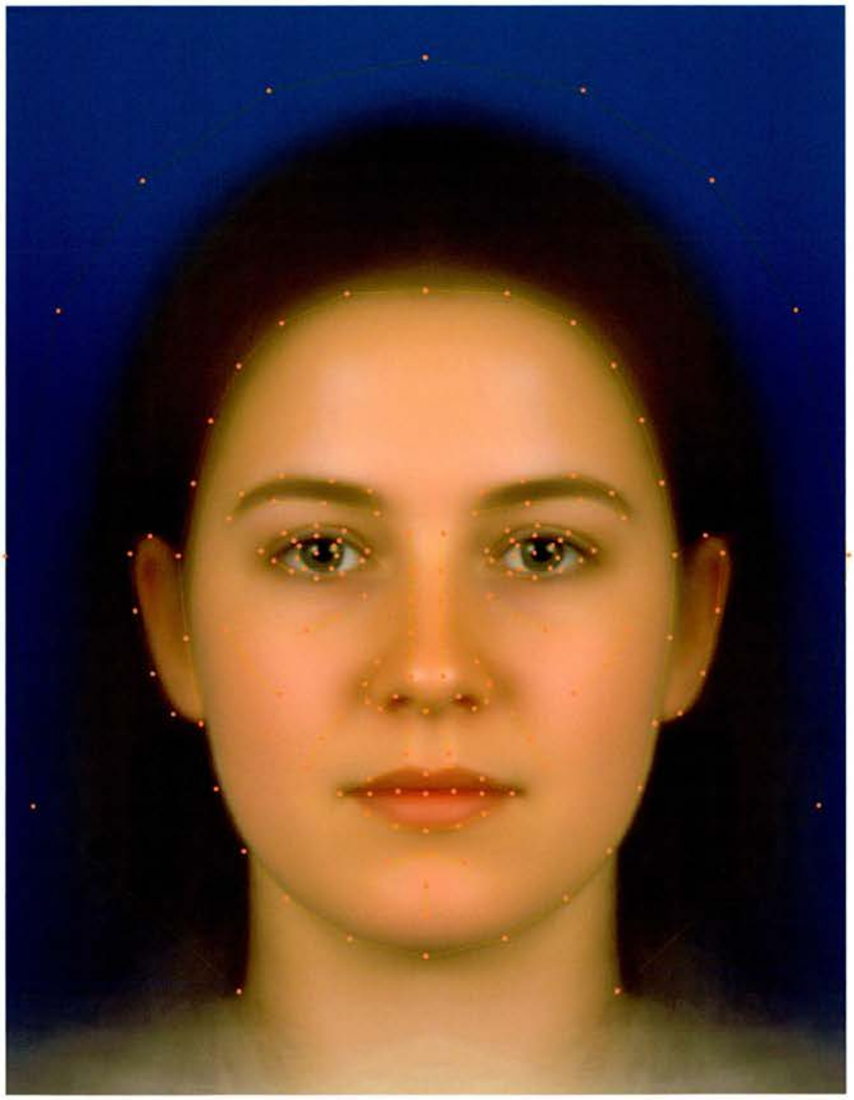


Figure 23: 174 delineation points used to mark facial features on a female average face

7.1.3 Calculating average colour

To generate average colour information to fill the average face shape (section 7.1.2), a two-stage process is used. Firstly, the colour information of each individual face is warped into the average face shape. To achieve this, the delineation data is used to tessellate the surface of the face into small triangles (Figure 24). Each one of these triangles can then be stretched to the same size and shape as the equivalent triangle in the average face shape. This process is analogous to putting the original face image on a thin rubber sheet, and then stretching or shrinking it into the shape of the average face (Figure 25). Secondly, the red, green and blue (RGB) colour values for each pixel location in the average shape are calculated by averaging the RGB values for the equivalent pixels in each of the individual warped faces.

The composite face of the set is now complete, and any consistent properties of the faces in the set should be extracted and manifested in the average (a collection of old male faces produces an average that looks like an old male, for example).

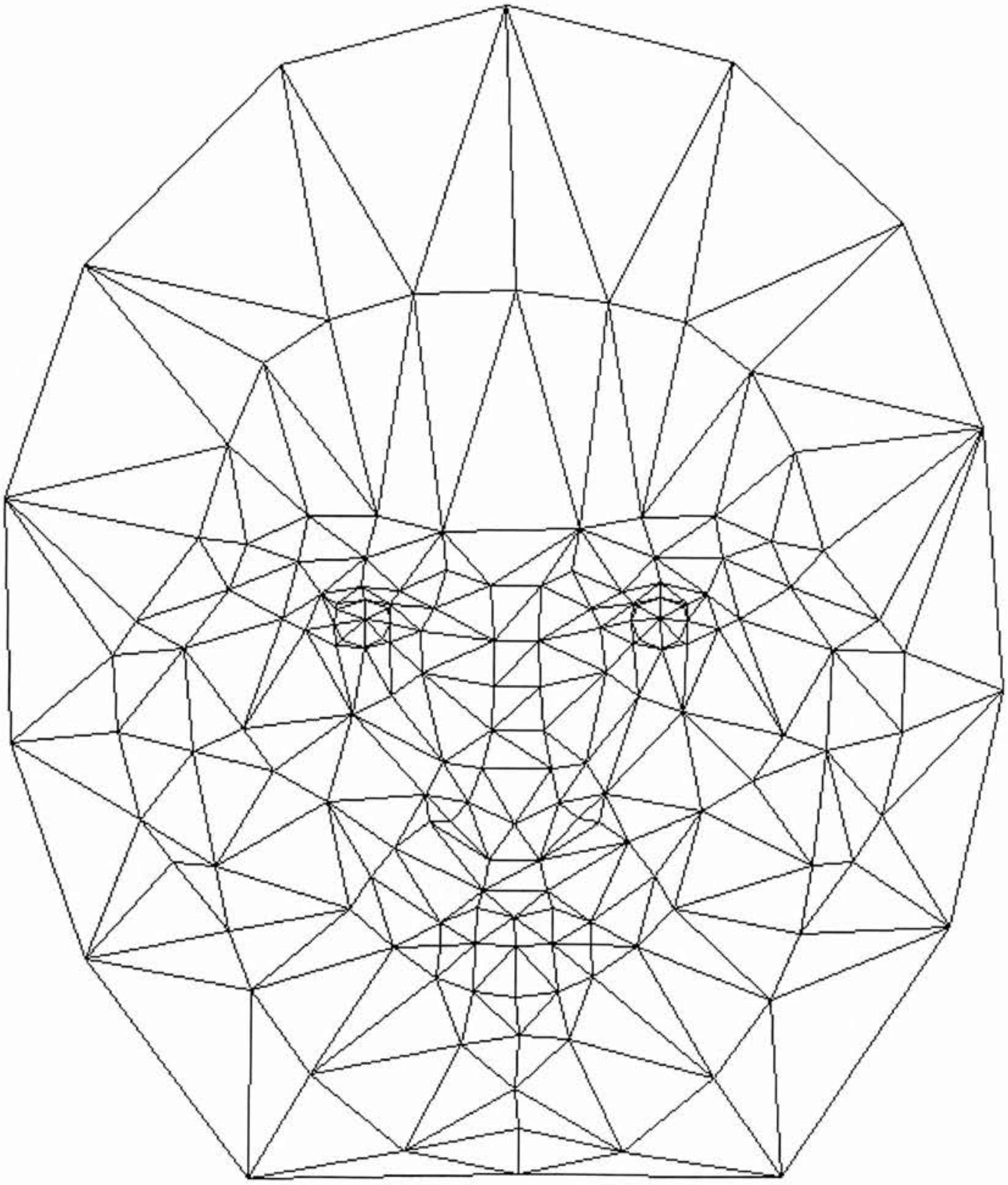


Figure 24: Triangles linking delineation points used for warping of individual faces into average shape.

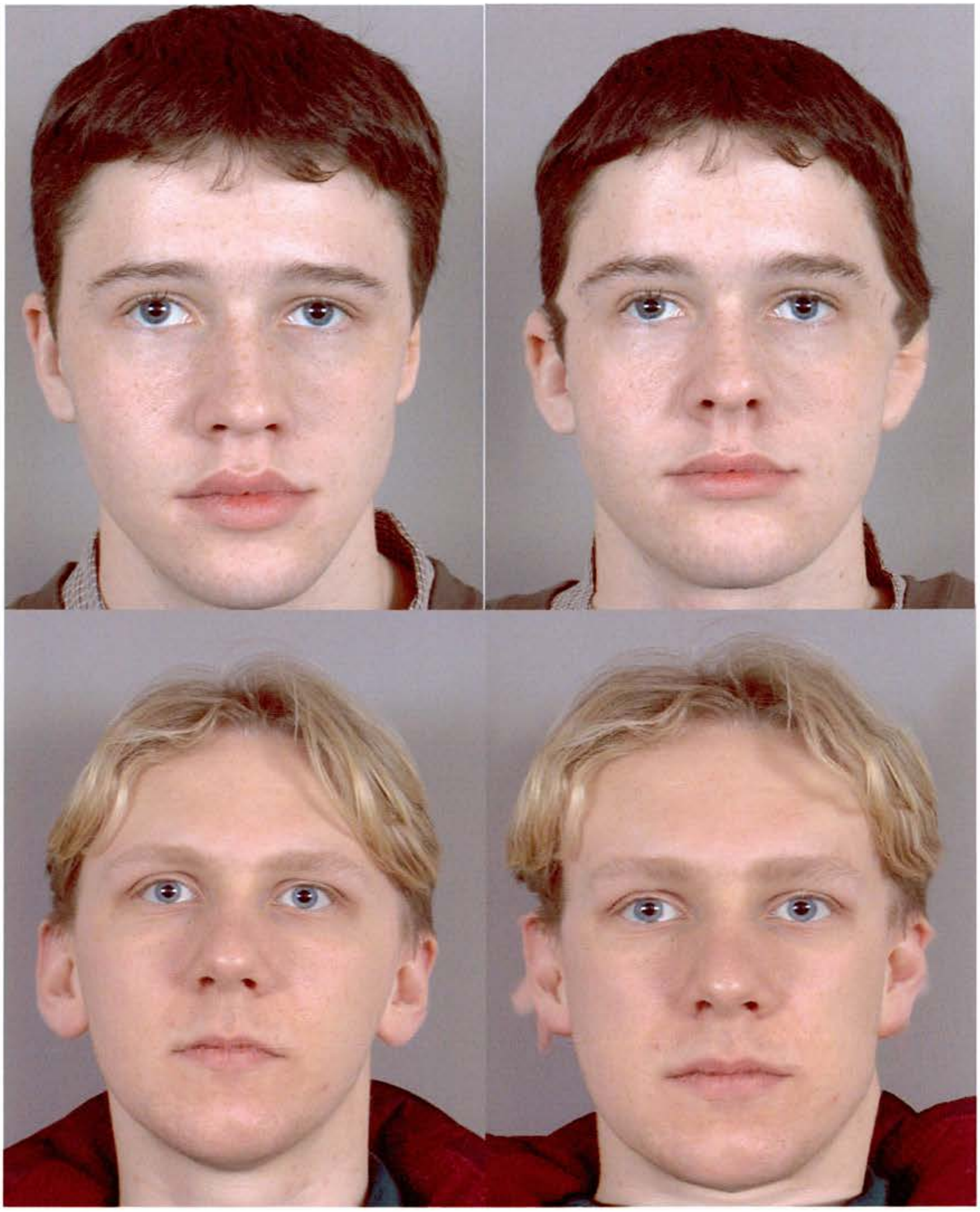


Figure 25: Two individuals (left, top and bottom) warped into an average male face shape (right, top & bottom). Once all individuals in the set have been rendered into the average shape, the mean RGB values of each pixel across the set are calculated to generate average colour information.

7.1.4.1 set 1: caucasian male and female (KJL).

25 British Caucasian males (age 19-23, mean 20.6) and 30 British Caucasian females (age 19-22, mean 20.6) at the University of St. Andrews were photographed using Fuji Chrome 100 ASA 35mm film in front of a standard blue background (to ensure colour constancy from film to film). Two flashguns (one left and one right of the sitter) provided diffuse lighting preventing shadows falling on sitters' faces. The images (each 2048*3072 pixels) were stored on Kodak Photo-CD, and delineated and manipulated at this size. Male and female composites were generated using the techniques described above, and are shown in Figure 26.

Figure 26: set KJL.

7.1.4.2 set 2: japanese male and female (KJL_JAP).

28 Japanese males (age 20-23, mean 21.6) and 28 Japanese females (age 20-22, mean 21.4) were photographed at Otemon-Gakuin University, Japan. Kodak 400 ASA 35mm film was used, with standard lighting provided by one flash mounted on the camera. Kodak Photo-CD Images were cropped from 2048*3072 pixels to 1961*2001 pixels to remove the background before delineation and manipulation. Male and female composites are shown in Figure 27.

Figure 27: set KJL_JAP.

7.1.4.3 set 3: caucasian male and female (IPV_TW)

40 female and 21 male Caucasian second year undergraduate students (Mean age 20 years) were photographed on 35mm film (Fuji Provia, 200 ASA). Subjects faced the camera two metres in front of a standard blue background. Two shots of each individual were taken. Diffused flash lighting from two flashguns was used to prevent shadows falling on the sitter's face. The films were developed and transferred to Kodak photo-CD (1024*1536 pixels). Kodak grey cards were used to ensure standardised hue, lightness and saturation between films during developing. Male and female composites are shown in Figure 28.

Figure 28: set IPV_TW.

7.1.4.4 *set 4: caucasian male and female (IPV_2)*

Conditions as set 3 above, but constructed from 18 males (mean age, 19.8 years) and 38 females (mean age, 20.8 years). Male and female composites are shown in Figure 29.

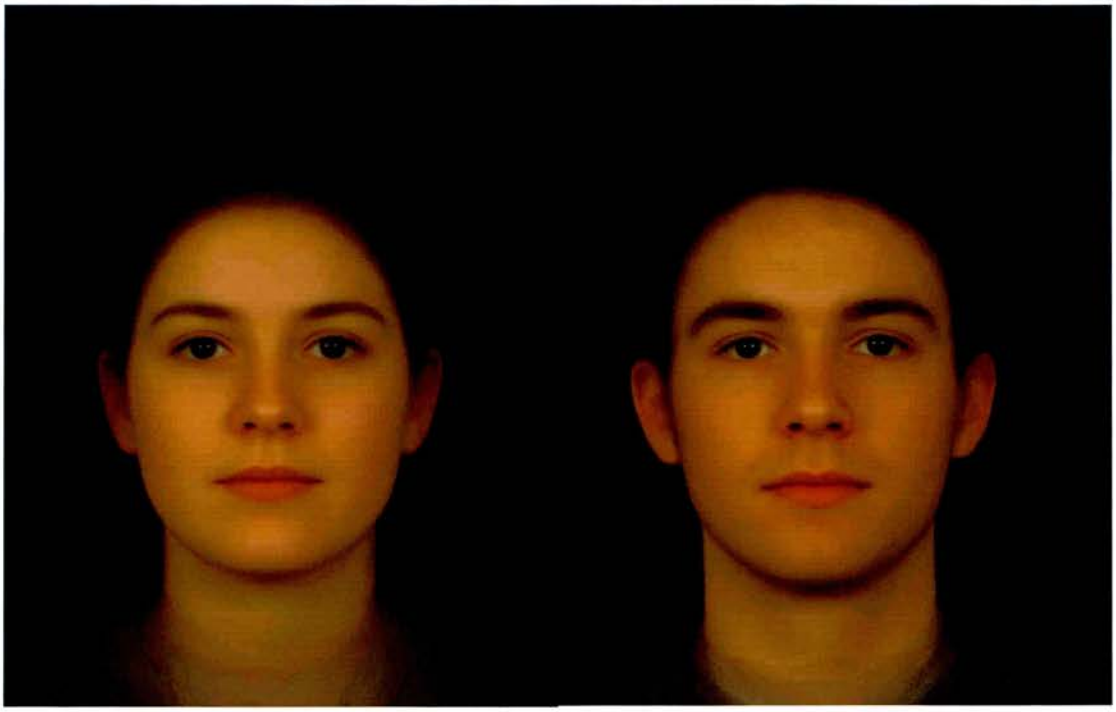


Figure 29: set IPV_2.

7.1.4.5 set 5: caucasian male and female (JEN)

A final pair of composites was constructed from 26 males (19.7 years) and 37 females (18.7 years). All were students at the University of St. Andrews. Sitters were photographed under diffuse flash lighting from two flashguns (as above) using a FUJIX DS-300 digital camera. Images were stored as Tagged Image Format (.tif) at 1000*1280 pixel resolution, before transfer to SGI workstations for delineation and manipulation. Male and female composites are shown in Figure 30.

Figure 30: set JEN.

7.2 Masking

Averaging techniques as described above do not generate convincing textures for hair; averaging across many individuals removes texture information resulting in a blurred hair area in composites. To eliminate hair cues, facial stimuli can be masked. The outline of the face is defined by straight lines connecting 26 of delineation points marking the hair and chin outline. The masks are made by removing the background around this outline and replacing it with a black mask. Blurring across the mask edge is necessary to avoid the image looking like a many sided polygon. This blurring is achieved by convolving the mask and a 2-D Gaussian function – effectively smoothing the edges. Masked and unmasked versions of stimuli from set IPV_TW are shown in figure 31.

Figure 31: unmasked and masked versions of the female IPV_TW.

7.3 Caricaturing: increasing and reducing sexual dimorphism

Caricaturing is the main technique used for manipulating sexual dimorphism in the composite stimuli used in the experiments reported here. Essentially, the process used to exaggerate of reduce levels of sexual dimorphism ('masculinity' and 'femininity') in faces is a development of digital caricaturing techniques (Brennan, 1985, Benson & Perrett, 1991). Exaggerating or reducing sex differences in face size and/or shape requires a precise measurement of what the differences are. Before the differences in face shape between male and female average faces are calculated, both face shapes must be aligned. Two alignment techniques are used in the stimuli presented in this thesis, detailed below.

7.3.1 Manipulating dimorphism in facial proportions *without* manipulating size dimorphism.

This technique first aligns male and female shapes by matching eye separations before calculating sex specific differences. As male faces are larger than female faces, the male prototype is scaled and rotated ('normalised') to equalise the inter-pupillary distance. Now the difference between each equivalent delineation point on the male and female shapes can be expressed as a two dimensional vector. To construct 'feminised' male face shapes, every feature point on the male face can be moved a prescribed distance along a vector toward its correspondent point on the aligned female average. 'Masculinised' face shapes can be created by exaggerating the vector differences between male and female faces. The colour information from the original male face shape is now rendered into the 'masculinised' or 'feminised' face shapes to complete the images.

Normalising opposite sex faces on both pupil positions introduces sizing artefacts into the stimuli. As the male average face must be shrunk to match the eye-separation of the female composite, any naturally occurring size dimorphism in male and female faces is lost. In addition, male eyes tend to be relatively widely set when compared to female eyes. Transforms based on stimuli aligned by eye separation actually *reverse* the apparent size differences in male and female faces: masculinising can make faces smaller (Figure 32)

7.3.2 Manipulating dimorphism in facial proportions *with* size dimorphism

To overcome the problems of aligning faces on both eye positions to generate male faces with exaggerated or reduced levels of shape dimorphism, male and female averages can be normalised on a point midway along a horizontal line between the eyes of both averages. This process requires that the faces be rotated so that the centres of the pupils are at the same level. Once this has been achieved and the faces normalised, any transform based on differences between the male and female shapes include size dimorphism (figure 32)

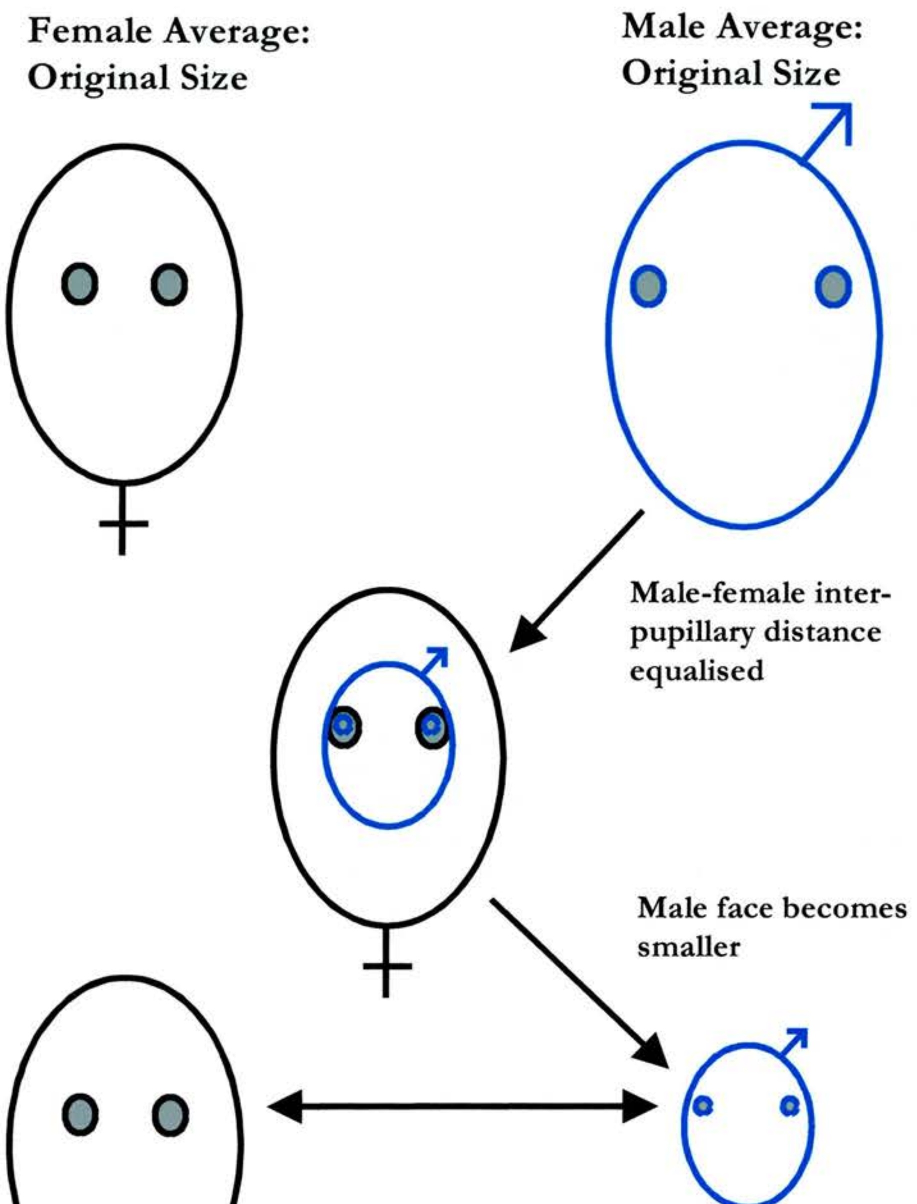


Figure 32: Normalisation without natural size dimorphism.

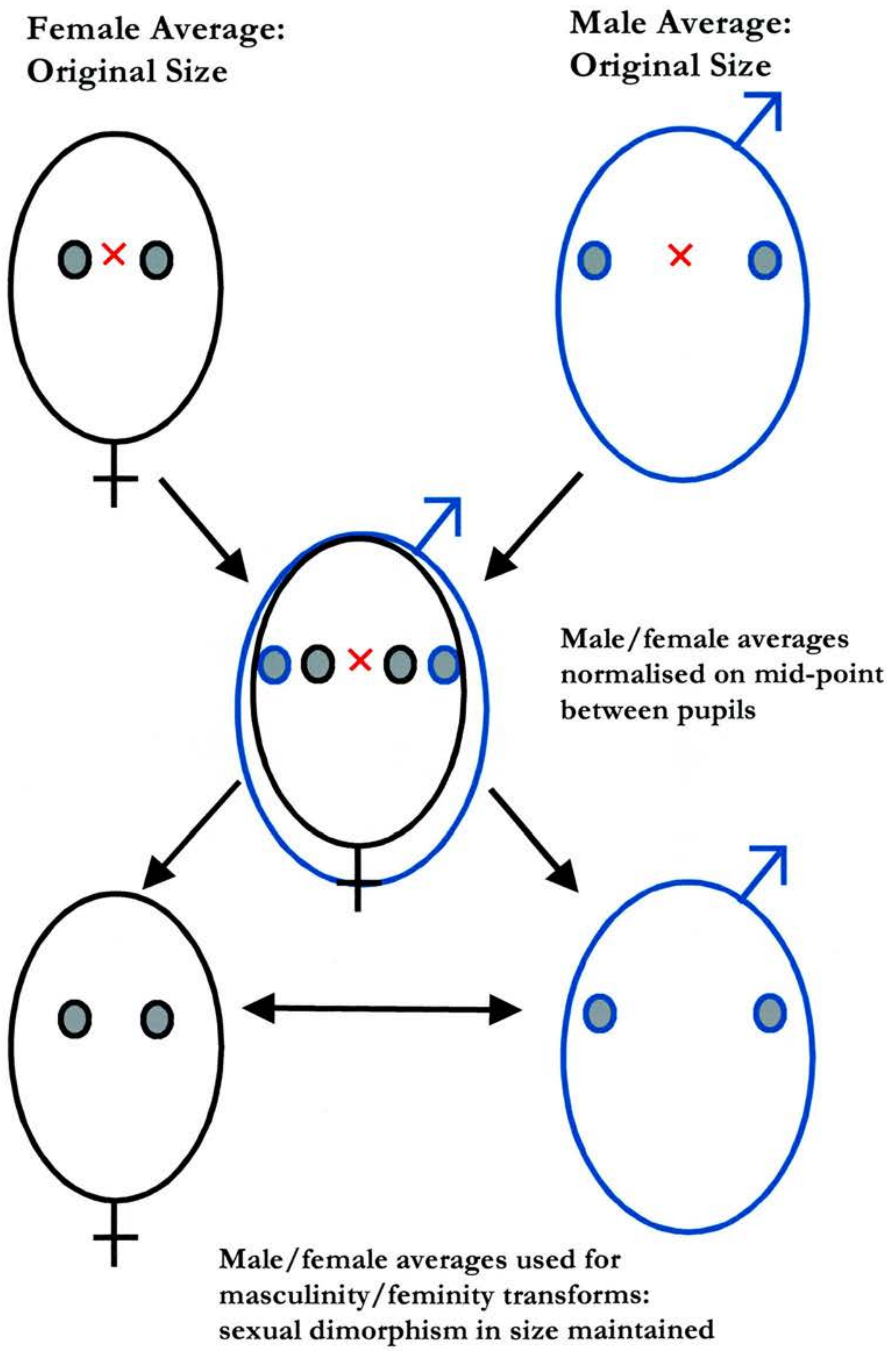
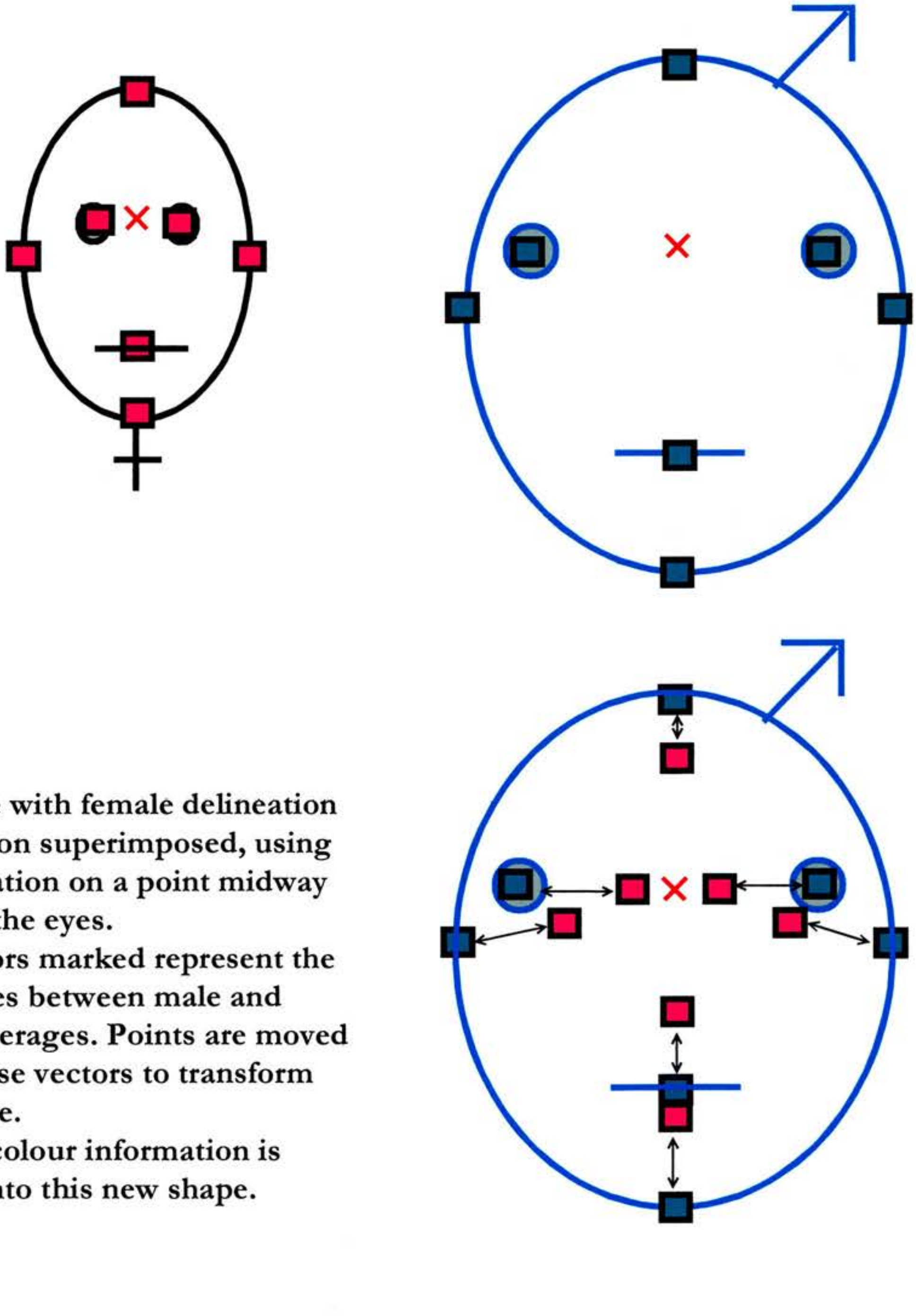


Figure 33: Normalisation maintaining natural size dimorphism.



Delineated Female Average:

Delineated Male Average:

Male face with female delineation information superimposed, using normalisation on a point midway between the eyes.
The vectors marked represent the differences between male and female averages. Points are moved along these vectors to transform face shape.
Average colour information is warped into this new shape.

Figure 34: Schematic illustration of male-female vector differences used in transforms.

7.4 Symmetry: giving composites perfect symmetry

Symmetry may influence attractiveness judgements (see section 2.3.1). When creating stimuli using caricaturing techniques, any small asymmetries present in the composites used for the transforms may be exaggerated. Keeping constant symmetry in stimuli prevents any possible confound. A symmetrical version of a composite shape can be generated by averaging the XY position (relative to a mid-line, perpendicular to and bisecting the inter-pupillary line) of corresponding pairs of delineation points on the left and right sides of the face. Symmetrical colour information is created by rendering each face that went into the composite and its mirror image into the new symmetrical shape.

7.5 Interactive testing

7.5.1 Interactive techniques on Silicon Graphics Workstations

Experiments reported here presented stimuli in one of two ways: either statically, (photographic quality printouts or still images on a computer monitor) or 'interactively'. Interactive techniques used here employ hardware graphics manipulations on Silicon Graphics Impact workstations that allow subjects to change the parameters of computer presented stimuli on-screen in real time, by moving the mouse left or right.

To generate an interactive continuum allowing a subject to manipulate the levels of masculinity or femininity in a face, the shape and colour information from the two endpoints is required. In the experiments reported here, the endpoints of the continua (normally a 50% masculinised and a 50% feminised version of the same composite) have identical colour information. Moving the mouse left or right alters the amount of the shape transform in the face displayed; for example, if the far left mouse position resulted in the 50% feminised image being displayed, and the far right mouse position would display the 50% masculinised postion, and intermediate mouse positions would display intermediate face shapes. Hardware morphing allows the colour information to be rendered into the shapes as they are presented on screen.

7.5.2 Interactive techniques for Internet Browsers.

To facilitate testing of participants and to explore the possibilities of large scale psychological testing using the world wide web, an interactive testing program was developed to run on any computer running Microsoft Internet Explorer (version 4 or later) or Netscape Navigator (release 4.0 or later). Instead of using real-time morphing (available on Silicon Graphics workstations) a series of static images are prepared. For example, to prepare a 50% feminised to a 50% maculinised continuum, 11 images in 10% steps were rendered on a Silicon Graphics workstation (50, 40, 30, 20, 10% feminised, the average, 50, 40, 30, 20, 10% masculinised). These images were transferred to a PC and converted to .gif format to reduce file size. A testing program was developed using Java that allowed a left-right mouse movement to change the image displayed: adjacent images in the continuum were displayed. This mimicked the effects of real time morphing. Although all the data presented in this thesis was collected from subjects in the laboratory, future work may be able to employ techniques such as this to test large numbers of subjects across the WWW.

8 General Discussion

Darwinian approaches to the study of facial attractiveness are based on the premise that an attractive face is a biological 'ornament' that signals valuable information to a potential mate. Evolutionary approaches to interpersonal attraction have grown in numbers over recent years, and have provided a fruitful theoretical base from which testable hypotheses have been generated and, in many cases, supported. One potential criticism of this field of research, however, is that it is almost exclusively structuralist in nature. Perception of facial attractiveness is assumed to be data-driven; the properties of a particular set of facial features are taken to be the same, irrespective of the perceiver. From this perspective, attractive faces are perceived as such independently of any of the characteristics of the observer, such as culture or upbringing. Systematic studies of the differences in facial attractiveness judgements between individuals, or of an individual's changing judgements of facial attractiveness over time, are rare. Although individual differences in attractiveness judgements obviously exist, such behavioural variation is rarely given a biological interpretation.

In fact, evolutionary biologists studying the sexual behaviour of other species provide a framework within which individual differences in human attractiveness judgements may be seen to have parallels with the behavioural variation shown by other species. Whilst human individuals appear to share some basic criteria of facial attractiveness, they may learn the 'fine grain' of the faces they find attractive in an analogous way to the learning or imprinting experiences of some other animals early in their lifetime. Furthermore, individuals may adopt different reproductive strategies as a result of life history factors and these differing strategies may lead to adaptive preferences for different face types. In this thesis, I have tried to examine various aspects of both 'structuralist' approaches to male facial attractiveness (experiments 1 & 2), and individual variation in female judgements of male attractiveness as a result of menstrual cycle status (experiments 3, 4 & 5) and self-similarity (experiments 6 & 7).

Whilst exaggerated femininity has been shown to be attractive in female faces, the role of sexual dimorphism in male facial attractiveness is more contentious. Experiments 1 & 2 propose that a key factor in male attractiveness is perceived personality. Female preferences for 'feminine' or 'baby' male faces can perhaps be explained in terms of the personality attributions that such faces attract. Although hypotheses based on indicator models of sexual selection suggest that masculinised faces should be preferred, such faces elicit negative personality attributions (coldness, dominance, and dishonesty, for example). Cross-culturally, personality factors are reported to be the most important factor in mate choice for both sexes (Buss, 1989). It seems inconceivable that personality attributions have no effect on attractiveness judgements. Personality attributions, though stereotypic, may predict behaviour. For example, ratings of perceived dishonesty from facial appearance correlate with the face owner's willingness to participate in deceptive behaviour (Bond, Berry & Omar, 1994). Preferences for feminised male faces may have been favoured by selection if they are linked to increased paternal investment in offspring, indicating sexual selection for direct phenotypic benefits rather than heritable immuno-competence.

Experiments 1 & 2, in common with most other research in the field, find that there is considerable agreement on which faces are attractive across individuals and cultures. Nonetheless, individual differences in attractiveness judgements certainly exist, as demonstrated both by the less than perfect agreement within and across experimental studies and by simple commonsense. The other experiments reported in this thesis examine several possible origins of such differences within a biological model of physical attractiveness.

Individual differences in women's preferences for sexual dimorphism in male faces may reflect complex, context-dependent decisions based on the probability of conception, the type of relationship sought and the current relationship status of an individual. Experiments 3, 4 and 5 indicate that menstrual cycle phase influences the relative attractiveness of sexual dimorphism in male faces. Both masculinity and femininity may be associated with behaviours that have costs and benefits to reproductive success (Perrett et al, 1998; Berry & Wero, 1983). Indicator models of sexual selection predict preferences for males with exaggerated sexual traits (i.e. masculine faces), which may signal potential benefits for offspring in terms of heritable fitness, but in humans also imply costs due to potentially decreased paternal investment. For example, Gangestad & Thornhill (1997) show that symmetrical men may be less likely to invest in a relationship, and have more extra-pair sex. Femininity in male faces may be associated with the opposite collection of characteristics, such as warmth, honesty and a willingness to invest in offspring (Perrett et al, 1998).

Females may subconsciously balance the costs and benefits of various reproductive strategies in the context of life history factors (e.g. presence/absence of a partner), preferring relatively feminine looking males when looking for long-term investment in offspring, or more masculine men when heritable benefits are more important. Alternatively, some women may adopt a mixed reproductive strategy and cuckold a primary partner whose relatively feminine appearance suggests co-operation in parental care by occasionally seeking extra-pair copulations with 'masculine' males when conception is most likely, and hence benefit without cost. Similar patterns of sexual activity are found in non-human species (Graves et al, 1993; Andersson, 1994). Individual differences in the faces that females find attractive may reflect strategies to improve reproductive success that vary according to circumstance: a factor that may lead to the difficulties researchers have encountered when trying to characterize attractiveness in male faces. The adoption of mixed reproductive strategies may be the result of psychological adaptations favoured by selection, but it is unclear at this point whether extra-pair paternity is common enough, or offers sufficient benefits, to drive selection for such behaviour in humans.

Experiments 6 and 7 examine the additional possibility that early childhood experience and learning may influence later sexual preferences, in a fashion somewhat analogous to 'imprinting' mechanisms common in birds and, in a weaker form, mammals. Although no positive evidence for such effects was found in these experiments, this may reflect weaknesses in design. Ongoing experiments are indicating that parental age may influence attractiveness judgements in a way that is analogous to imprinting – offspring of older parents are more tolerant of age in potential partners.

8.1 Conclusions

Attractive faces appear to have some stable qualities that generalize across cultures and are salient to individuals very early in infancy, indicating that normative social information alone is unlikely to be responsible for attractiveness judgements.

People nonetheless differ in the faces that they find attractive. Tentatively, I have suggested here that at least some of this behavioural variation may reflect biologically adaptive processes similar to those documented in other species. Changes within individuals over time may reflect short-term changes in hormonal state (as demonstrated by cyclic changes in face preferences in females) or changes in circumstance such as the presence or absence of a partner. Although self-similarity does not seem to be a great influence on facial attractiveness judgements, imprinting on other family characteristics (e.g. parental age) may yet be shown to contribute to differences in attractiveness judgements between individuals.

9 References

Adams, D., A. Burt, et al. (1978). "Rise in female-initiated sexual activity at ovulation and its suppression by oral contraception." New England Journal of Medicine **299**: 1145-1150.

Aiello, L. and C. Dean (1990). An introduction to human evolutionary anatomy. London, Academic Press.

Albright, L., D. A. Kenny, et al. (1988). "Consensus in personality judgments at zero acquaintance." Journal of Personality and Social Psychology **55**(3): 387-395.

Albright, L., T. E. Malloy, et al. (1997). "Cross-cultural consensus in personality judgments." Journal of Personality and Social Psychology **72**(3): 558-569.

Alexander, R. D., J. L. Hooglund, et al. (1979). Sexual dimorphisms and breeding systems in pinnipeds, ungulates, primates, and humans. Evolutionary Biology and Human Social Behavior. N. Chagnon and W. Irons. MA, Duxbury Press.

Alexander, R. D. and K. M. Noonan (1979). Concealment of ovulation, parental care and human social evolution. Evolutioary Biology and Human Social Behavior. W. N. Chagnon & Irons. MA, Duxbury Press.

Alley, T. R. (1988). Physiognomy and social perception. Social and Applied Aspects of Perceiving Faces. T. R. Alley. Hillsdale, NJ, Erlbaum.

Alley, T. R. and M. R. Cunningham (1991). "Averaged faces are attractive, but very attractive faces are not average." Psychological Science **2**(2): 123-125.

Alley, T. R. and K. A. Hildebrandt (1988). Determinants and consequences of facial aesthetics. Social and Applied Aspects of Perceiving Faces. T. R. Alley. Hillsdale, NJ, Erlbaum.

Allison, D. B., M. C. Neale, et al. (1996). "Assortative mating for relative weight: Genetic implications." Behavior Genetics **26**(2): 103-111.

Allport, G. W. and H. S. Odbert (1936). "Trait names: A psycholexical study." Psychological Monographs **47**.

Ambady, N., M. Hallahan, et al. (1995). "On judging and being judged accurately in zero-acquaintance situations." Journal of Personality and Social Psychology **69**(3): 518-529.

Andersson, M. (1994). Sexual Selection. Princeton, NJ, Princeton University Press.

Baker, A. H., A. C. Denning, et al. (1998). "How accurate are women's estimates of date of onset of next menses?" Psychology and Health **13**(5): 897-908.

Baker, R. R. and M. A. Bellis (1995). Human Sperm Competition: copulation, masturbation and infidelity. London, Chapman & Hall.

Bancroft, J., D. Sanders, et al. (1983). "Mood, sexuality, hormones, and the menstrual cycle. 3. Sexuality and the role of androgens." Psychosomatic Medicine **45**: 509-516.

Barrett, L. F. and P. R. Pietromonaco (1997). "Accuracy of the five-factor model in predicting perceptions of daily social interactions." Personality and Social Psychology Bulletin **23**: 1173-1187.

Bateson, P. (1978). "Sexual imprinting and optimal outbreeding." Nature **273**: 659-660.

Bateson, P. (1980). "Optimal Outbreeding and the development of Sexual Preferences in Japanese Quail." Z. Tierpsychol. **53**: 231-244.

Bateson, P. (1983). Optimal Outbreeding. Mate Choice. P. Bateson. Cambridge, Cambridge University Press: 257-277.

Bateson, P. (1988). Preferences for close relations in Japanese Quail. Acta XIX Congressus Internationalis Ornithologici. H. Ouellet. Ottawa, Universtity of Ottawa Press. **1:** 961-972.

Bean, J. A., J. D. Leeper, et al. (1979). "Variations in the reporting of menstrual histories." American Journal of Epidemiology **109**: 181-185.

Benshoof, L. and R. Thornhill (1979). "The evolution of monogamy and the loss of estrous in humans." J. Soc. Biol. Struc. **2**: 95-106.

Benson, P. and D. Perrett (1992). "Face to face with the perfect image." New Scientist **133**: 32-35.

Benson, P. J. and D. I. Perrett (1991). "Perception and recognition of photographic quality facial caricatures: Implications for the recognition of natural images." European Journal of Psychology **3**: 105-135.

Berland, K. J. H. (1993). "Reading character in the face: Lavater, Socrates & Physiognomy." Word & Image **9**: 252-269.

Bernstein, I. H., T. Lin, et al. (1982). "Cross- vs. within-racial judgments of attractiveness." Perception and Psychophysics **32**(6): 495-503.

Berry, D. (1991). "Attractive faces are not all created equal: Joint effects of facial babyishness and attractiveness on social perception." Personality and social psychology bulletin **17**(5): 523-533.

Berry, D. S. and S. Brownlow (1989). "Were the physiognomists right? Personality correlates of facial babyishness." Personality and Social Psychology Bulletin **15**: 266-279.

Berry, D. S. and J. C. Landry (1997). "Facial maturity and daily social interaction." Journal of Personality and Social Psychology **72**: 570-580.

Berry, D. S. and L. Z. McArthur (1985). "Some components and consequences of a babyface." Journal of Personality and Social Psychology **48**(2): 312-323.

Berry, D. S. and J. L. F. Wero (1993). "Accuracy in face perception: A view from ecological psychology." Ecological Psychology: 497-519.

Berry, D. S. and L. Zebrowitz-McArthur (1988). The imoact of age-related craniofacial changes on social perception. Social and Applied Aspects of Perceiving Faces. T. R. Alley. Hillsdale, NJ, Erlbaum.

Berscheid, E., K. Dion, et al. (1971). "Physical attractiveness and dating choice: A test of the matching hypothesis." Journal of Experimental and Social Psychogy **7**: 173-189.

Bettelheim, B. (1969). The children of the dream. New York, Macmillan.

Blaffer Hrdy, S. (1979). "The evolution of human sexuality: the latest word and the last." The Quarterly Review of Biology **54**: 309-314.

Blaffer Hrdy, S. and P. Whitten (1987). Patterning of sexual activity. Primate Societies. B. B. Smutts, D. L. Cheney, R. M. Seyfarth, R. Wrangham and T. T. Strusaker. Chicago, University of Chicago Press: 370-384.

Bond, C. F., D. S. Berry, et al. (1994). "The kernal of truth in judgements of deceptiveness." Basic and Applied Social Psychology **15**: 523-534.

Booth, A. and J. Dabbs (1993). "Testosterone and men's marriages." Social Forces **72**: 463-477.

Borkenau, P. and A. Liebler (1992). "Trait inferences: Sources of validity at zero acquaintance." Journal of Personality and Social Psychology **62**(4): 645-657.

Borod, J. C., H. S. Caron, et al. (1981). "Asymmetry of facial expression related to handedness, footedness, and eyedness: a quantitative study." Cortex **17**: 381-390.

Borod, J. C. and E. Koff (1983). "Hemiface mobility and facial expression asymmetry." Cortex **19**: 327-332.

Botwin, M. D., D. M. Buss, et al. (1997). "Personality and mate preferences: Five factors in Mate selection and marital satisfaction." Journal of Personality **65**(1): 107-136.

Bouchard, T. J. and M. McGue (1981). "Familial Studies of Intelligence; A Review." Science **212**: 1055-1059.

Brace, C. L., K. R. Rosenberg, et al. (1987). "Gradual change in human tooth size in the late pleistocene and post-pleistocene." Evolution **41**: 705-720.

Brandt, A. (1980). "Face reading: The persistence of physiognomy." Psychology Today(Dec.): 90-96.

Brennan, S. E. (1985). "The caricature generator." Leonardo **18**: 170-178.

Burley, N. (1979). "The evolution of concealed ovulation." American Naturalist **114**: 835-858.

Burley, N. (1983). "The meaning of assortative mating." Ethology and Sociobiology **4**: 191-203.

Burling, R. (1985). "Garo marriage to the wife's daughter: further support for the Westermarck hypothesis for the origin of the incest taboo." American Anthropologist. **87**: 130-133.

Burt, A. (1992). "'Concealed ovulation' and sexual signals in primates." Folia Primatologica **58**: 1-6.

Burt, D. M. and D. I. Perrett (1995). "Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and colour information." Proceedings of the Royal Society of London, B **259**: 137-143.

Buss, D. M. (1985). "Human mate selection." American Scientist **73**: 47-51.

Buss, D. M. (1989). "Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures." Behavioral and Brain Sciences **12**: 1-49.

Buss, D. M. (1991). "Do women have evolved mate preferences for men with resources? A reply to Smuts." Ethology and Sociobiology **12**: 401-408.

Buss, D. M. (1999). Evolutionary psychology: the new science of the mind. Boston, Allyn & Bacon.

Buss, D. M. and M. Barnes (1986). "Preferences in human mate selection." Journal of Personality and Social Psychology **50**: 559-570.

Buss, D. M. and D. P. Schmitt (1993). "Sexual Strategies Theory: An evolutionary perspective on human mating." Psychological Review **100**(2): 204-232.

Butlin, R. K. and T. Tregenza (1997). "Is speciation no accident?" Nature **387**: 551-553.

Calcagno, J. M. and K. R. Gibson (1988). "Human dental reduction - natural selection or the probable mutation effect." American Journal of Physical Anthropology **1988**(77): 505-515.

Calcagno, J. M. and K. R. Gibson (1989). "The increasing population-densities effect - a new, but once again improbable, mechanism of human dental reduction." American Journal of Physical Anthropology **78**: 201.

Cattell, R. B. (1966). The Scientific Analysis of Personality. Chicago, Aldine.

Cavalli-Sforza, L. L. and W. F. Bodmer (1971). The genetics of human populations. San Francisco., W.H. Freeman & Co.

Chabbert-Buffert, N., C. Djakoure, et al. (1998). "Regulation of the human menstrual cycle." Frontiers in Neuroendocrinology **19**: 151-186.

Christenfeld, N. and E. A. Hill (1995). "Whose baby are you?" Nature **378**: 669.

Clark, P. J. and J. N. Spuhler (1959). "Differential fertility in relation to body dimensions." Human Biology **31**(2): 121-137.

Cleeton, G. U. and F. B. Knight (1924). "Validity of character judgement based on external criteria." Journal of Applied Psychology **8**: 215-231.

Cunningham, M. R. (1986). "Measuring the physical in physical attractiveness: Quasi-experiments on the sociobiology of female facial beauty." Journal of Personality and Social Psychology **50**(5): 925-935.

Cunningham, M. R., A. P. Barbee, et al. (1990). "What do women want? Facialmetric assessment of multiple motives in the perception of male facial physical attractiveness." Journal of Personality and Social Psychology **59**(1): 61-72.

Cunningham, M. R., A. R. Roberts, et al. (1995). ""Their ideas of beauty are, on the whole, the same as ours": Consistency and variability in the cross-cultural perception of female attractiveness." Journal of Personality and Social Psychology **68**(2): 261-279.

Daly, M. and M. Wilson (1983). Sex, Evolution & Behavior. Boston, PWS.

Daly, M. and M. I. Wilson (1999). "Human evolutionary psychology and animal behaviour." Animal Behaviour **57**: 509-519.

Darwin, C. (1871). The descent of man, and selection in relation to sex. London, John Murray.

Dawkins, R. (1976). The Selfish Gene. Oxford, Oxford University Press.

Dawkins, R. (1979). "Twelve misunderstandings of kin selection." Zeitschrift fur Tierpsychologie **51**: 184-200.

De Haan, M., M. H. Johnson, et al. (in review). "Recognition of individual faces and average face prototypes by 1- and 3-month-old infants." Developmental Psychology.

Deutsch, F. M., C. M. Zalenski, et al. (1986). "Is there a double standard of aging?" Journal of Applied Social Psychology **16**(9): 771-785.

Dixson, A. F. (1983). "Observations on the evolution and behavioral significance of 'sexual skin' in female primates." Advances in the Study of Behavior **13**: 63-106.

Doty, R. L., P. J. Snyder, et al. (1981). "Endocrine, cardiovascular, and psychological correlates of olfactory sensitivity changes during the human menstrual cycle." Journal of Comparative and Physiological Psychology **95**: 45-60.

D'Udine, B. and E. Alleva (1983). Early experience and sexual preferences in rodents. Mate Choice. P. Bateson. Cambridge, Cambridge University Press.

Dugatkin, L. A. and J. J. Godin (1998). "How females choose their mates." Scientific American(4): 46-51.

Enquist, M. and S. Ghirlanda (1998). "Evolutionary Biology - the secrets of faces." Nature **394**: 826-827.

Epstein, E. and R. Guttman (1982). "Mate selection in man: Evidence, Theory, and Outcome." Social Biology **31**(3-4): 243-276.

Eysenck, H. J. (1953). The Structure of Human Personality. New York, Wiley.

Feingold, A. (1992). "Good-looking people are not what we think." Psychological Bulletin **111**(2): 304-341.

Fisher, R. A. (1915). "The evolution of sexual preference." Eugenics Review **7**: 184-192.

Fisher, R. A. (1930). The Genetical Theory of Natural Selection. Oxford, Clarendon Press.

Foley, R. E. and P. C. Lee (1991). "Ecology and energetics of encephalization in hominid evolution." Philosophical Transcripts of the Royal Society of London, B **334**: 223-232.

Folstad, I. and A. J. Karter (1992). "Parasites, bright males and the immunocompetence handicap." American Naturalist **139**: 603-622.

Ford, C. S. and F. A. Beach (1951). Patterns of Sexual Behaviour. London, Eyres & Spottiswood.

Frost, P. (1994). "Preference for darker faces in photographs at different phases of the menstrual cycle: preliminary assessment of evidence for a hormonal relationship." Perceptual and Motor Skills **79**: 507-514.

Gage, A. R., J. T. Manning, et al. (submitted). "Testosterone, cortisol and jaw size in men." .

Gallagher, J. (1976). "Sexual imprinting: effects of various regimes of social experience on mate preference in Japanese quail *Coturnix coturnix japonica*." Behaviour **57**: 91-114.

Galton, F. J. (1878). "Composite Portraits." Anthropological Institute of Great Britain and Ireland **8**: 132-142.

Gangestad, S. W. and D. M. Buss (1993). "Pathogen prevalence and human mate preferences." Ethology and Sociobiology **14**: 89-96.

Gangestad, S. W. and R. Thornhill (1997). "The evolutionary psychology of extrapair sex: The role of fluctuating asymmetry." Evolution and Human Behaviour **18**(2): 69-88.

Gangestad, S. W. and R. Thornhill (1998). "Menstrual Cycle Variation in Women's Preferences for the scent of symmetrical men." Proceedings of the Royal Society of London, B **265**: 927-933.

Gangestad, S. W., R. Thornhill, et al. (1994). "Facial attractiveness, developmental stability and fluctuating asymmetry." Ethology and Sociobiology **15**: 73-85.

Gibson, R. M. (1990). "Relationships between blood parasites, mating success and phenotypic cues in male sage grouse." American Zoologist **30**: 271-278.

Grammer, K. (1993). "5-alpha-androst-16en-3alpha-on: a male pheromone? A brief report." Ethology and Sociobiology **14**: 201-208.

Grammer, K. and R. Thornhill (1994). "Human (*homo sapiens*) facial attractiveness and sexual selection: the role of symmetry and averageness." Journal of Comparative Psychology **108**: 233-242.

Graves, J., J. Ortega-Ruano, et al. (1993). "Extra-pair copulations and paternity and shags: do females choose better males?" Proceedings of the Royal Society of London, B **253**: 3-7.

Graves, R., H. Goodglass, et al. (1982). "Mouth asymmetry during spontaneous speech." Neuropsychologia **20**: 371-381.

Graves, R. and T. Landis (1990). "Asymmetry in mouth opening during different speech tasks." International Journal of Psychology **25**: 179-189.

Greenwood, P. J. (1980). "Mating systems, philopatry and dispersal in birds and mammals." Animal Behaviour **28**: 1140-62.

Griffiths, R. W. and P. R. Kunz (1973). "Assortative Mating: A study of physiognomic homogamy." Soc. Bio. **20**(4): 448-453.

Grossman, C. J. (1985). "Interactions between the gonadal steroids and the immune system." Science **227**: 257-261.

Haaq, C. M. and P. K. Donahoe (1998). "Regulation of Sexual Dimorphism in Mammals." Physiological Reviews **78**(1): 1-33.

Hamilton, W. D. (1964). "The genetical evolution of social behaviour, 1 & 2." Journal of Theoretical Biology **7**: 1-52.

Hamilton, W. D. (1980). "Sex versus non-sex versus parasite." Oikos **35**: 282-290.

Hamilton, W. D. (1990). "Mate Choice Near or Far." American Zoologist **30**: 341-352.

Hamilton, W. D. and M. Zuk (1982). "Heritable true fitness and bright birds: A role for parasites?" Science **218**: 384-387.

Harcourt, A. H., P. H. Harvey, et al. (1981). "Testis weight, body weight and breeding system in primates." Nature **293**: 55-57.

Hawkes, K., J. F. O'Connell, et al. (1991). "Hunting income patterns among the Hadza: big game, common goods, foraging goals and the evolution of the human diet." Phil. Trans. R. Soc. Lond. B **334**: 243-251.

Heywood, J. S. (1989). "Sexual selection by the handicap mechanism." Evolution **43**: 1387-1397.

Hill, C. T., Z. Rubin, et al. (1976). "Breakups before marriage: The end of 103 affairs." Journal of Social Issues **32**: 147-168.

Hillgarth, N. (1990). "Parasites and female choice in the ring-necked pheasant." American Zoologist **30**: 227-233.

Hillgarth, N. and J. C. Wingfield (1997). Testosterone and Immunosuppression in vertebrates: implications for parasite mediated sexual selection. Parasites and Pathogens. N. E. Beckage. New York, Chapman &Hall.

Hinsz, V. B. (1989). "Facial resemblance in engaged and married couples." J. Soc. Pers. Relationships **6**: 223-229.

Holmes, W. G. and P. W. Sherman (1983). "Kin Recognition in Animals." American Scientist. **71**(Jan-Feb): 46-55.

Hrdy, S. B. (1981). The woman that never evolved. Cambridge, Ma., Harvard University Press.

Hume, D. K. and R. Montgomerie (1999). Facial attractiveness and symmetry in humans. Human Behavior and Evolution Society, Utah.

Immelman, K. (1975). "Ecological significance of imprinting and early learning." Annual review of Ecology and Systematics **6**: 15-37.

Isaac, G. (1978). "The food-sharing behavior of protohuman hominids." Scientific American **238**: 90-108.

Jackson, L. A. (1992). Physical appearance and Gender: Sociobiology and sociocultural perspectives. Albany, State University of New York Press.

James, W. H. (1983). "Correlation between father's age and husband's age: a case of imprinting." J. Biosoc. Sci. **15**: 115-117.

Jensen, A. R. (1978). Genetic and behavioral effects of non-random mating. Human Variation: Biogenetics of age, race and sex. C. E. Noble, Osborne, R.T., & Weyle, N. New York, Academic Press.

Jochle, W. (1973). "Coitus-induced ovulation." Contraception **7**(6): 523-564.

Johnson, R. C., F. M. Ahern, et al. (1980). "Secular change in degree of assortative mating for ability?" Behavioral Genetics. **10**: 1-8.

Johnston, V. S. and M. Franklin (1993). "Is beauty in the eye of the beholder?" Ethology and Sociobiology **14**: 183-199.

Johnstone, R. A. (1994). "Female preference for symmetrical males as a by-product of selection for mate recognition." Nature **372**: 172-175.

Jones, D. (1995). "Sexual selection, physical attractiveness, and facial neoteny." Current Anthropology **36**(5): 723-748.

Jones, D. and K. Hill (1993). "Criteria of facial attractiveness in five populations." Human Nature **4**(3): 271-296.

Kalick, S. M., L. A. Zebrowitz, et al. (1998). "Does human facial attractiveness honestly advertise health? Longitudinal data on an evolutionary question." Psychological Science **9**(1): 8-13.

Keating, C. F. (1985). "Gender and the physiognomy of dominance and attractiveness." Social Psychology Quarterly **48**: 61-70.

Keating, C. F., A. Mazur, et al. (1981). "A cross-cultural exploration of physiognomic traits of dominance and happiness." Ethology and Sociobiology **2**: 41-48.

Kendrick, K. M., M. R. Hinton, et al. (1998). "Mothers determine male sexual preferences." Nature **395**: 229-230.

Kenny, D. A., L. Albright, et al. (1994). "Consensus in interpersonal perception: acquaintance and the big five." Psychological Bulletin **116**: 245-258.

Kenny, D. A., C. Horner, et al. (1992). "Consensus at zero acquaintance: replication, behavioral cues and stability." Journal of Personality and Social Psychology **62**: 88-97.

Kenrick, D. T. and R. C. Keeve (1992). "Age preferences in mates reflect sex differences in human reproductive strategies." Behavioral and Brain Sciences **15**: 75-133.

Kirkpatrick, M. (1982). "Sexual selection and the evolution of female choice." Evolution **36**: 1-12.

Koeslag, J. and P. Koeslag (1994). "Koinophilia." Jouranl of Theoretical Biology **167**: 55-65.

Koeslag, J. H. (1994). "Koinophilia replaces random mating in populations subject to mutations with randomly varying fitnesses." J. Theor. Biol. **171**: 341-345.

Kowner, R. (1996). "Facial asymmetry and attractiveness judgment in developmental perspective." Journal of Experimental Psychology: Human Perception and Performance. **22**(3): 662-675.

Kurland, J. A. (1979). Paternity, mother's brother, and human sociality. Evolutionary Biology and Human Social Behavior. N. C. W. Irons. North Sciute, MA, Duxbury.

Lancaster, J. (1975). Primate Behavior and the Emergence of Human Culture. New York, Holt, Reinhart & Winston.

Lande, R. (1981). "Models of speciation by sexual selection on polygenic traits." Proceedings of the National Academy of Science, USA **78**: 3721-3725.

Langlois, J. H., J. M. Ritter, et al. (1991). "Facial diversity and infant preferences for attractive faces." Developmental Psychology **27**(1): 79-84.

Langlois, J. H. and L. A. Roggman (1990). "Attractive faces are only average." Psychological Science **1**(1): 115-121.

Langlois, J. H., L. A. Roggman, et al. (1987). "Infant preferences for attractive faces: Rudiments of a stereotype?" Developmental Psychology **23**(3): 363-369.

Langlois, J. H., L. A. Roggman, et al. (1994). "What is average and what is not average about attractive faces?" Psychological Science **5**: 214-219.

Lee, K. J. (1998). Computer graphic studies of facial characteristics. School of Psychology. St. Andrews, University of St. Andrews.

Lickliter, R. and G. Gottlieb (1985). "Social interaction with siblings is necessary for visual imprinting of species-specific maternal preferences in ducklings." J. Comp. Psy. **99**(4): 371-379.

Liggett, J. (1974). The Human Face. London, Constable.

Lockwood, C. A., B. G. Richmond, et al. (1996). "Randomization procedures and sexual dimorphism in *Australopithecus afarensis*." Jouranl of Human Evolution **31**: 537-548.

Lorenz, K. (1943). "The innate forms of potential experience." Zietschrift fur Tierpsychologie **5**: 234-409.

Low, B. S. (1990). "Marriage systems and pathogen stress in human societies." American Zoologist **30**: 325-339.

MacArthur, R. H. (1962). "Some generalized theorems of natural selection." Proceedings of the National Academy of Sciences (USA) **48**: 1893-1897.

Macintyre, S. and A. Sooman (1991). "Non-paternity and prenatal genetic screening." The Lancet **338**: 869.

Malina, R. M., H. A. Selby, et al. (1983). "Assortative mating for phenotypic characteristics in a Zapotec community in Oaxaca, Mexico." Journal of Biosocial Science. **15**: 273-280.

Manning, J. T. and A. T. Chamberlain (1993). "Fluctuating asymmetry, sexual selection and canine teeth in primates." Proceedings of the Royal Society of London, B. **251**: 83-87.

Maynard Smith, J. (1978). The Evolution of Sex. Cambridge, Cambridge University Press.

Mazur, A. and A. Booth (1998). "Testosterone & dominance in men." BBS **21**: 353-371.

Mazur, A., C. Halpern, et al. (1994). "Dominant looking males copulate earlier." Ethology and Sociobiology **15**: 87-94.

Mazur, A., E. J. Susman, et al. (1997). "Sex difference in testosterone response to a video game contest." Evolution and Human Behavior **18**: 317-326.

McArthur, L. A., K. (1983-1984). "Impressions of baby-faced adults." Social Cognition **2**(4): 315-342.

McArthur, L. Z. and D. S. Berry (1987). "Cross-cultural agreement in perceptions of babyfaced adults." Journal of Cross-cultural Psychology **18**(2): 165-192.

McCabe, J. (1983). "FDB Marriage: Further support for the Westermarck hypothesis of the incest taboo?" American Anthropologist **85**: 50-69.

McCrae, R. R. and P. T. Costa (1987). "Validation of the Five Factor model of personality across instruments and observers." Journal of Personality and Social Psychology **52**(1): 81-90.

Mealey, L., R. Bridgestock, et al. (1999). "Symmetry and perceived facial attractiveness." Journal of Personality and Social Psychology **76**: 151-158.

Merikangas, K. R. (1982). Arch. Gen. Psych. **39**: 1173-1180.

Meyer, D. A. and M. W. Quong (1999). "The bio-logic of facial geometry." Nature **397**: 661-662.

Miller, G. F. (1997). A review of sexual selection and human evolution: How mate choice shaped human nature. Evolution and human behavior: ideas, issues and applications. C. C. D. Krebs, Lawrence Erlbaum.

Möller, A. P. (1992). "Female swallow preference for symmetrical male sexual ornaments." Nature **357**: 238-240.

Möller, A. P. (1997). "Developmental stability and fitness: A review." American Naturalist **149**: 916-942.

Möller, A. P. and J. Hoglund (1991). "Patterns of fluctuating asymmetry in avian feather ornaments: Implications for models of sexual selection." Proceedings of the Royal Society of London, B **245**: 1-5.

Möller, A. P. and A. Pomiankowski (1993). "Fluctuating asymmetry and sexual selection." Genetica **89**: 267-279.

Möller, A. P., M. Soler, et al. (1995). "Breast asymmetry, sexual selection, and human reproductive success." Ethology and Sociobiology **16**: 207-219.

Möller, A. P. and R. Thornhill (1997). "Developmental Stability is Heritable." Journal of Evolutionary Biology **10**: 69-76.

Möller, A. P. and R. Thornhill (1997). "A meta-analysis of the heritability of developmental stabillity." Journal of Evolutionary Biology **10**: 1-16.

Möller, A. P. and R. Thornhill (1998). "Bilateral Symmetry and Sexual Selection: A meta-analysis." American Naturalist **151**: 174-192.

Morris, D. (1994). The Naked Ape, Vintage.

Mueller, U. and A. Mazur (1997). "Facial dominance in *Homo sapiens* as honest signaling of male quality." Behavioral Ecology **8**(5): 569-579.

Murdock, G. P. (1967). Ethnogrpahic Atlas. Pittsburgh, University of Pittsburgh Press.

Murray, R. D. S., E. O. (1983). "The role of dominance and intrafamilial bonding in the avoidance of close inbreeding." J. Human Evoution **12**: 481-486.

Norman, W. T. (1963). "Toward an adequate taxonomy of personlity attributes: Replicated factor structure in peer nominated personality ratings." Journal of Abnormal and Social Psychology **66**: 574-583.

Norris, K. (1993). "Heritable variation in a plumage indicator of viability in male great tits *parsus major*." Nature **362**: 537-539.

notgot

Ober, C., H. T., et al. (1998). "Leukocyte antigen matching and fetal loss - results of a ten year prospective study." Human Reproduction **13**: 33-38.

O'Connell, J. F., K. Hawkes, et al. (1999). "Grandmothering and the evoltuion of *Homo erectus*." Journal of Human Evolution **36**: 461-485.

O'Donald, P. (1980). Genetic Models of Sexual Selection. Cambridge, Cambridge University Press.

Owens, I. P. F., & Short, R.V. (1995). "Hormonal basis of sexual dimorphism in birds: implications for new theories of sexual selection." TREE **10**(1): 44-47.

Packer, C. (1979). "Inter-troop transfer and inbreeding avoidance in *Papio Anubis*." Animal Behaviour **27**: 1-36.

Pagel, M. (1997). "Desperately concealing father: a theory of parent-infant resemblance." Anim. Behav. **53**: 973-981.

Parker, G. A., R. R. Baker, et al. (1972). "The origin and evolution of gametic dimorphism and the male-female phenomenom." Journal of Theoretical Biology **36**: 529-553.

Parsons, P. A. (1990). "Fluctuating asymmetry: an epigenetic measure of stress." Biological Review **65**: 131-145.

Partridge, L. (1983). Non-random mating and offspring fitness. Mate Choice. P. Bateson. Cambridge, Cambridge University Press.

Passini, F. T. N., W.T. (1966). "A universal conception of personallity structure?" Journal of Personlity and Social Psychology **4**: 44-49.

Pawlowski, B. and R. I. M. Dunbar (1999). "Impact of market value on human mate choice." Proceedings of the royal Society of London, B. **266**: 281-285.

Penton-Voak, I., D. Perrett, et al. (in press). "Computer graphic studies of the role of facial similarity in attractiveness judgements." Current Psychology.

Perrett, D. I., May, K.A., & Yoshikawa, S. (1994). "Facial shape and judgements of female attractiveness." Nature **368**: 239-242.

Perrett, D. I., D. M. Burt, et al. (in press). "Symmetry and Human Facial Attractiveness." Evolution and Human Behavior.

Perrett, D. I., K. J. Lee, et al. (1998). "Effects of sexual dimorphism on facial attractiveness." Nature **394**: 884-887.

Perusse, D. (1993). "Cultural and reproductive success in industrial societies: Testing the relationship at proximate and ultimate levels." Behavioral and Brain Sciences **16**: 267-322.

Petrie, M. (1994). "Improved growth and survival of offspring of peacocks with more elaborate trains." Nature **371**: 598-599.

Pianka, E. R. (1970). "On r- and K-selection." American Naturalist **104**: 592-597.

Pollard, J., J. Shepherd, et al. (1999). "Average faces are average faces." Current Psychology **18**(1): 98-103.

Price, T. D., D. Schluter, et al. (1993). "Sexual selecion when the female directly benefits." Biol. J. Linn. Soc. **48**: 187-211.

Queller, D. C. (1997). "Why do females care more than males?" Proceedings of the Royal Society of London, B. **264**: 1555-1557.

Rabin, I. A. (1965). Growing up in a kibbutz. New York, Springer.

Raven, P. and G. Johnson (1988). Understanding Biology. St. Louis, Times Mirror/Mosby College Publishing.

Regan, P. C. (1996). "Rhythms of desire:the association between menstrual cycle phases and female sexual desire." The Canadian Journal of Human Sexuality **5**(3): 145-156.

Rhodes, G., & Tremewan, T. (1996). "Averageness, exaggeration and facial attractiveness." Psychological Science **7**(2): 105-110.

Rhodes, G., Byatt, G., Hickford, C. & Jeffery, L. (review). "Sex-typicality and facial attractiveness: Are supermale and superfemale faces super-attractive?" British Journal of Psychology.

Rhodes, G., F. Proffitt, et al. (1998). "Facial symmetry and the perception of beauty." Psychonomic Bulletin Review **5**: 659-669.

Ridley, M. (1993). The Red Queen. London, Viking.

Roberts, D. F. (1977). "Assortative mating in man: Husband/Wife Correlations in Physical Characteristics." Supplement to the Bulletin of The Eugenics Society. **2**: Whole Part.

Rowland, D. A., & Perrett, D.I. (1995). "Manipulating facial appearance through shape and color." IEEE Computer Graphics and Applications **15**(5): 70-76.

Ruff, C. B., E. Trinkaus, et al. (1997). "Body mass and encephalization in Pleistocene Homo." Nature **387**: 173-176.

Ruff, C. B., E. Trinkaus, et al. (1993). "Postcranial robusticity in homo. 1. Temporal trends and mechanical interpretation." American Journal of Physical Anthropology **91**: 21-53.

Rushton, J. P. (1988). "Genetic similarity, mate choice, and fecundity in humans." Ethology and Sociobiology **9**: 329-333.

Rushton, J. P. (1989). "Genetic similarity, human altruism, and group selection." Behavioral and Brain Sciences **12**: 503-559.

Rushton, J. P. (1995). "Genetic similarity theory and human assortative mating: a reply to Russell and Wells." Anim. Behav. **50**: 547-549.

Rushton, J. P. and I. R. Nicholson (1988). "Genetic similarity theory, intelligence, and human mate choice." Ethology and Sociobiology **9**: 45-57.

Russell, R. J. H., & Wells, P.A. (1994). "Human assortative mating: more questions concerning genetic similarity theory." Anim. Behav. **47**: 463-464.

Russell, R. J. H., & Wells, P.A. (1995). "Human assortative mating: more questions concerning genetic similarity theory." Anim. Behav. **50**: 550-553.

Ryan, M. J. and S. Rand (1990). "The sensory basis of sexual selection for complex calls in the Tungara frog, *physalaemus pustulosus* (sexual selection for sensory exploitation)." Evolution **44**: 305-314.

Samuels, C. A., Butterworth, G., Roberts, T., Graupner, L., & Hole, G. (1994). "Facial aesthetics: babies prefer attractiveness to symmetry." Perception **23**: 823-831.

Scheib, J. E., S. W. Gangestad, et al. (in press). "Facial attractiveness, symmetry, and cues to good genes." Proceedings of the Royal Society of London, B.

Schröder, I. (1993). "Concealed ovulation and clandestine copulation: a female contribution to human evolution." Ethology and Sociobiology **14**: 381-389.

Sciulli, P. W. and M. C. Mahaney (1991). "Phenotypic Evolution in prehistoric Ohio Amerindians - Natural selection versus random genetic drift in tooth size reduction." Human Biology **63**: 499-511.

Seemanova, E. (1971). "A study of the children of incestuous matings." Human Heredity **21**: 108-128.

Seger, J. (1985). "Unifying genetic models for the evolution of female choice." Evolution **39**: 1185-1193.

Shepher, J. (1971). "Mate selection among second generation Kibbutz adolescents and adults: Incest avoidance and negative imprinting." Archives of Sexual Behavior **1**(4): 293-307.

Silber, M. (1994). "Menstrual cycle and work schedule: effects of women's sexuality." Archives of Sexual Behavior **23**: 397-404.

Sillen-Tullberg, B. and A. P. Möller (1993). "The relationship between concealed ovulation and mating systems in anthropoid primates: A phylogenetic analysis." American Naturalist **141**(1): 1-25.

Singh, D. (1993). "Adaptive significance of female physical attractiveness: Role of waist-to-hip ratio." Journal of Personality and Social Psychology **65**(2): 293-307.

Singh, D. (1993). "Body shape and women's attractiveness - the critical role of waist-to-hip ratio." Human Nature **4**(3): 297-321.

Singh, D. (1994). "Waist-to-hip ratio and judgment of attractiveness and healthiness of female figures by male and female physicians." International Journal of Obesity **18**: 731-737.

Singh, D. (1995). "Female health, attractiveness and desirability for relationships: Role of breast asymmetry and waist to hip ratio." Ethology and Sociobiology **16**: 465-481.

Skinner, M. and B. Mullen (1991). "Facial asymmetry in emotional expression: A meta-analysis of research." British Journal of Social Psychology **30**: 113-124.

Slater, A., C. Von der Schulenberg, et al. (1998). "Newborn babies prefer attractive faces." Infant Behaviour and Development **21**(2): 345-354.

Sloman, S. S., L. (1988). "Mate selection in the service of human evolution." J. Social biol. Struct. **11**: 457-468.

Small, M. F. (1993). Female Choice: Sexual Behavior of Female Primates. Ithaca, Cornell University Press.

Small, M. F. (1996). ""Revealed" ovulation in humans?" Journal of Human Evolution **30**: 483-488.

Smith, R. H. (1979). "On selection for inbreeding in polygynous animals." Heredity **43**: 205-211.

Spiro, M. E. (1965). Children of the Kibbutz. New York, Schocken Books.

Spuhler, J. N. (1968). "Assortative mating with respect to physical characteristics." Eugenics Quarterly **15**(2): 128-140.

Stanislaw, H. and F. J. Rice (1988). "Correaltion between sexual desire and menstrual cycle characteristics." Archives of Sexual Behavior **17**: 499-508.

Stewart, D. E. (1989). "Positive changes in the pre-menstrual period." Acta Psychiatra Scandinavia **79**: 400-405.

Susanne, C. (1977). "Heritability of anthropological characters." Human Biology **49**(4): 573-580.

Swaddle, J. P. & Cuthill, I.C. (1995). "Asymmetry and human facial attractiveness: symmetry may not always be beautiful." Proceedings of the Royal Society of London **261**: 111-116.

Symons, D. (1979). The evolution of human sexuality. Oxford, Oxford University Press.

Symons, D. (1994). Beauty is in the adaptions of the beholder: The evolutionary psychology of human female sexual attractiveness. Sexual Nature/ Sexual Culture. P. R. P. Abramson, S.D. Chicago, University of Chicago Press.

Talmon, Y. (1964). "Mate selection in collective settlements." American Sociological Review **29**: 491-508.

Ten Cate, C. B., P. (1989). "Sexual Imprinting and a preference for 'Supernormal' Partners in Japanese Quail." Animal Behaviour **38**(2): 356-357.

Thiessen, D. and B. Gregg (1980). "Human assortative mating and genetic equilibrium: An evolutionary perspective." Ethology and Sociobiology **1**: 111-140.

Thornhill, R., & Gangestad, S.W. (1993). "Human Facial Beauty: Averageness, Symmetry, and Parasite Resistance." Human Nature **4**(3): 237-269.

Thornhill, R., & Gangestad, S. (1996). "The evolution of human sexuality." TREE **11**(2): 98-102.

Thornhill, R. and S. W. Gangestad (in press). "The scent of symmetry: a human sex pheromone that signals fitness?" Evolution and Human Behavior.

Thornhill, R., S. W. Gangestad, et al. (1995). "Human female orgasm and mate fluctuating asymmetry." Animal Behavior **50**: 1601-1615.

Thornhill, R. G., S.W. (1994). "Human fluctuating asymmetry and sexual behaviour." Psychological Science **5**: 297-302.

Thornhill, R. G., S.W. (in preparation). "Facial Attractiveness." Trends in Cognitive Science.

Tooby, J. (1982). "Pathogens, polymorphism and the evolution of sex." Journal of Theoretical Biology **97**: 557-576.

Tovee, M. J., Reinhardt, S., Emery, J.L., & Cornelissen, P.L. (1998). " Optimum body-mass index and maximum sexual attractiveness." Lancet **352**: 548.

Townsend, J. M., & Roberts, L.W. (1993). "Gender differences in mate preference among law students: Divergence and convergence of criteria." Journal of Psychology **127**(5): 507-528.

Townsend, J. M., Kline, J. & Wasserman, T.H. (1995). "Low investment copulation: sex differences in motivations and emotional reactions." Ethology & Sociobiology **16**: 25-51.

Townsend, J. M. W., T. (1998). "Sexual attractiveness: sex differences in assessment and criteria." Evolution and Human Behavior **19**: 171-191.

Trivers, R. L. (1972). Parental investment and sexual selection. Sexual selection and the descent of man, 1871-1971. B. Campbell. Chicago, Aldine: 136-179.

Udry, J. R. and N. Morris (1968). "Distribution of coitus in the menstrual cycle." Nature **200**: 593-596.

Urdy, R. J. and N. M. Morris (1977). "The distribution of events in the human menstrual cycle." Journal of Reproduction and Fertility **51**: 419-425.

Van Den Berghe, P. L. (1983). "Human inbreeding avoidance: culture in nature." Behavioral and Brain Sciences **6**: 91-123.

Van Den Berghe, P. L. and P. Frost (1986). "Skin color preference, sexual dimorphism and sexual selection: a case of gene culture co-evolution?" Ethnic and Racial Studies **9**(1): 87-113.

Vandenberg, S. G. (1972). "Assortative Mating or Who Marries Whom?" Behavior Genetics **2**(2/3): 127-157.

Vos, D. R. (1994). "Sex recognition in zebra finch males results form early experience." Behaviour **128**: 1-14.

Vos, D. R. (1995). "The role of sexual imprinting for sex recognition in zebra finches: a difference between males and females." Anim. Behav. **50**: 645-653.

Vos, D. R. (1995). "Sexual imprinting in Zebra-finch females: Do females develop a preference for males that look like their Father?" Ethology **99**: 252-262.

Waldman, B. (1987). "Mechanisms of kin recognition." Journal of Theoretical Biology **128**: 159-185.

Walster, E., Aronson, V. Abrahams, D. & Rottman, L. (1966). "Importance of physical attractiveness in dating behavior." Journal of Personality and Social Psychology **4**(5): 508-516.

Watson, D. (1989). "Strangers' ratings of the five robust personality factors: Evidence of a surpising convergence with self-report." Journal of Personality and Social Psychology **57**: 120-128.

Wedekind, C. (1992). "Detailed information about parasites as revealed by sexual ornamentation." Proceedings of the Royal Society of London, B **147**: 169-174.

Wedekind, C., & Folstad, I. (1994). "Adaptive or nonadaptive immunosuppression by sex hormones?" American Naturalist **143**(5): 936-938.

Wedekind, C., Seebeck, T., Bettens, F. & Paepke, A.J. (1995). "MHC-dependent mate preferences in humans." Proceedings of the Royal Society of London, B. **260**: 245-249.

Wedekind, C. and S. Furi (1997). "Body odour preferences in men and women: do they aim for specific MHC combinations or simply heterozygosity?" Proceedings of the Royal Society of London, B **264**: 1471-1479.

Westermarck, E. (1894). The history of human marriage. London, Macmillan.

Wiederman, M. W. D., S.L. (1998). "Evolution and sex differences in preferences for short-term mates: Results from a plociy capturing study." Evolution and Human Behavior **19**: 153-170.

Wilkinson, G. S. and P. R. Reillo (1994). "Female choice response to artificial selection on an exaggerated mael trait in a stalk eyed fly." Proceedings of the Royal Society of London, B **255**: 1-6.

Wilson, D. S. (1997). "Evolution - human groups as units of selection." Science **276**: 1816-1817.

Wilson, G. D., & Barrett, P.T. (1987). "Parental Characteristics and partner choice: Some evidence for Oedipal imprinting." J. Biosoc. Sci. **19**: 157-161.

Wolf, A. P. (1966). "Childhood association, sexual attraction and the incest taboo: A Chinese Case." Am. Anthropologist **68**: 883-898.

Wolf, A. P. (1993). "Westermarck Redivivus." Annu. Rev. Anthropol. **22**: 157-175.

Wolf, A. P. and C. S. Huang (1980). Marriage and adoption in China, 1845-1945. Stanford, Stanford University Press.

Wolf, M. E. and M. A. Goodale (1987). "Oral asymmetries during verbal and non-verbal movements of the mouth." Neuropsychologia **26**: 823-832.

Worthman, C. (1978). "Psychoendocrine study of human behavior: some interactions of steroid hormones with affect and behavior in !Kung San", Ph.D. thesis, Harvard University.

Wright, S. (1933). The roles of mutation, inbreeeding, crossbreeding and selection in evolution. Proceedings of the VIth International Conference on Genetics., Brooklyn Botanic Garden: New York.

Wylie, D. R. and M. A. Goodale (1998). "Left-sided oral asymmetries in spontaneous but not posed smiles." Neuropsychologia **25**: 823-832.

Yu, D. W. and G. H. Shepard (1998). "Is beauty in the eye of the beholder?" Nature **396**: 321-322.

Zahavi, A. (1975). "Mate selection: A selection for a handicap." Journal of Theoretical Biology **53**: 205-214.

Zajonc, R. B., Adelmann, P.K., Murphy, S.T. & Niendenthal, P.M. (1987). "Convergence in the Physical Appearance of Spouses." Motivation and Emotion **11**(4): 335-346.

Zei, G., P. Astofli, et al. (1981). "Correlation between father's age and husband's age: A case of imprinting?" J. Biosoc. Sci. **13**: 409-418.

Zei, G., P. Astolfi, et al. (1983). "Correlation between father's age and husband's age: A case of imprinting." Journal of Biosocial Science **15**: 116-117.

Zuk, M. (1992). "The role of parasites in sexual selection: Current evidence and future directions." Advances in the Study of Behavior **21**: 39-67.