The Resemblance of One-year-old Infants to Their Fathers: Refuting Christenfeld & Hill (1995)

To appear in the *Proceedings of the 22nd Annual Cognitive Science Society Conference*, (2000), NJ: LEA

Robert M. French, Serge Brédart, Johanne Huart, Christophe Labiouse
Department of Psychology
University of Liège, 4000 Liège, Belgium
rfrench; Serge.Bredart; jhuart; clabiouse @ulg.ac.be

Abstract

In 1995 Christenfeld and Hill published a paper that purported to show at one year of age, infants resemble their fathers more than their mothers. Evolution, they argued, would have produced this result since it would ensure male parental resources, since the paternity of the infant would no longer be in doubt. We believe this result is false. We present the results of two experiments (and mention a third) which are very far from replicating Christenfeld and Hill's data. In addition, we provide an evolutionary explanation as to why evolution would *not* have favored the result reported by Christenfeld and Hill.

Introduction

Science overwhelmingly favors positive results. To appreciate this, one need look no further than the almost exclusive emphasis on Type I error detection in the social sciences and the quasi-religious status of the inequality "$p < 0.05$." Since everyone knows that the null hypothesis cannot be proved, only rejected, it follows that the only results worth pursuing are those that involve the rejection of a null hypothesis. This is why it is so much harder to establish (and publish) negative results. But sometimes null hypotheses are rejected incorrectly and subsequently setting the record straight proves to be very difficult indeed. For every published rejection of a null hypothesis, a far greater number of failures-to-replicate are usually necessary to convincingly establish that the published result was most probably in error.

Although examples of this problem abound, it is instructive to briefly recall the well-known experiments on the chemical transfer of memory done by McConnell (1962) and others. Planarian worms were trained to respond in a certain "correct" way to a light source. These worms were then killed and their RNA fed to a new set of worms, who, by dint of having ingested the previous worms' RNA, would supposedly respond correctly to the light source more often than worms in a control group. Once this result, buttressed by theoretical arguments about the role of RNA in memory, was published it became very hard to unseat it, in spite of numerous failures-to-replicate (e.g., Bennett & Calvin, 1964; Byrne et al., 1966). Thus, even though numerous failure-to-replicate papers had begun to appear as early as 1964, many courses on memory into the 1970's still included McConnell's results on the chemical transfer of memory (see, for example, Munn, Fernald, & Fernald, 1969; Hilgard, Atkinson, & Atkinson, 1971).

In this paper we will present two negative results that we hope will help serve to establish the falsehood of a published result — namely, the claim of greater resemblance between one-year-old infants and their fathers than their mothers (Christenfeld and Hill, 1995). This result received very wide international attention when it was published in 1995. The result is now cited often but we believe, both for theoretical and empirical reasons, that it is wrong. In the present paper, we will present our own failures-to-replicate the original results and will give a theoretical justification for our results. We hope that this will lead other researchers to also critically examine the originally published results of Christenfeld and Hill before they become firmly, and in our opinion wrongly, entrenched as fact.

The remainder of this paper is organized as follows. We begin by briefly presenting the claim of Christenfeld and Hill (1995). This will be followed by the results of two independent experiments (i.e., different subjects, different sets of stimuli, etc.) that fail to replicate their results. We will then give a theoretical justification for why evolution would most likely have produced our results and not those of Christenfeld and Hill.

Christenfeld & Hill (1995): One-year old Infants Resemble Their Fathers

Christenfeld and Hill (1995) reported a result in 1995 that appeared in *Nature* and received considerable attention throughout the world, both in the scientific and the popular press. They claimed to have found greater facial resemblance between one-year-old children and their fathers than between one-year-old children and their mothers. Their result had wide appeal, in particular, because it seemed to agree with a prediction of evolutionary psychology (Gaulin and Schegel, 1980) — namely, that "It could then be to a baby's advantage to look like the father, to encourage paternal investment [on the part of the male parent]" (Christenfeld & Hill, 1995) since a mother can be quite sure that the baby is hers but the father cannot.

According to Christenfeld and Hill, greater father-child resemblance would be to the baby's advantage because it would encourage the father's investment in its survival, since he would be able to clearly identify

the child as his own. This would tend to produce a differential survival rate among children who, at age one (when they were most in need of resources from the father for their survival), looked like their fathers and those who did not.

Overview of the Two Experiments

In an initial experiment (not reported here) involving 200 subjects done soon after Christenfeld and Hill's paper first appeared, we were unable to reproduce their results. We thought that perhaps there might be some problem with the photographic stimuli we were using. (Christenfeld and Hill declined our request to make their original stimuli available.) We therefore created a second set of stimuli, careful to make sure that there the photos displayed no beards, glasses, hats, or other features that might distract from the identification task. However, once again, we failed to replicate Christenfeld and Hill's results. These results are reported in Experiment 1 (see Brédart & French, 1999).

We then created another, entirely new set of stimuli and designed the experiment to record participants' reaction times during identification. As before, we found virtually no difference in the level of correct identification of children and their real mother compared to children and their real father. Further, in addition to repeating the results of the first experiment, there was no significant difference between correct child-mother and child-father identification times.

We have now attempted to reproduce Christenfeld and Hill's result with three different sets of stimuli with three different groups of participants using two different measures (% of correct identification and reaction time). In no case did we find any significant differences in father-infant and mother-infant identification. In other words, we have what we believe to be good empirical evidence that belies the originally reported findings of Christenfeld and Hill.

Experiment 1

Subjects

One hundred and eighty undergraduate students at the University of Liège participated in the experiment. Thirty subjects (15 female and 15 male) were randomly assigned to each condition. Their ages ranged from 18 to 30 years (mean age = 21.84).

Stimuli and materials

Twenty-eight Caucasian families provided five photographs: three photographs of the same child at one year, three years, and five years, as well as one photograph of the mother and one photograph of the father taken when the child was approximately one year old. For fourteen families, the child was a girl, for the other fourteen families the child was a boy. The stimuli presented to subjects were scanned versions of these photographs (size = 5x4 cm) of faces. None of the faces had glasses, beards or moustaches.

Procedure

On each trial, participants were presented with the face of a child and, according to the condition, the faces of three women or three men. Their task was to identify the child's parent among the three presented adult faces. There were 28 trials (14 different girls and 14 different boys). The photographs were displayed in the same way as in the Christenfeld and Hill study: the child's face was presented in an upper position and the three adults' faces were placed beneath the child's face. The presentation positions of the adult photos were appropriately randomized. Participants were tested individually. Each were each presented with the 28 sets including one child and three possible parents in a different random order.

Results

The design of the experiment was as follows. The age of the child (one-year-old, three-year-old and five-year-old) and the sex of the parent were between-subjects factors while the sex of the child was a within-subjects factor. A 3 (age of the child) X 2 (sex of the parent) X 2 (sex of the child) ANOVA with repeated measures on the last factor revealed a significant main effect of the age of the child ($F(2,174) = 6.614$, $p < .01$), no main effect of the sex of the parent ($F(1,174) < 1$) and no main effect of the sex of the child ($F(1,174) < 1$). The analysis revealed no significant interaction between the first two factors ($F(2, 174) < 1$), no significant interaction between the second and the third factor ($F(1,174) < 1$) and no three-way interaction ($F(2,174) < 1$). The main effect of the age of the child was qualified by a significant interaction between this factor and the sex of the child ($F(2,174) = 5.988$, $p < .01$), but the magnitude of this interaction effect is low ($\eta^2 = 0.06$). This interaction was analyzed using Tukey HSD post-hoc tests. These tests showed that, while the level of parent identification from pictures of girls did not change across the three ages, it did for pictures of boys. Parent identification was better for five-year-old boys than from one-year-old boys ($p < .0001$) and than for three-year-old boys ($p < .01$). No significant difference appeared between one-year-old and three-year-old boys ($p = 0.56$). Post-hoc tests indicated no significant effect of the sex of the child on parent identification at age one, three or five (all p's > .10).

A control analysis taking the items as the random factor was also carried out. This analysis did not reveal any significant main effect of the sex of the child ($F(1,26) < 1$), of the sex of the parent ($F(1, 26) < 1$) and of the age of the child ($F(2, 52) = 1.982$, $p = 0.15$). Nor did it reveal any interaction effect (all p's > .20). The results of this control analysis confirmed that the significant interaction effect obtained in the preceding analysis was not a strong effect.

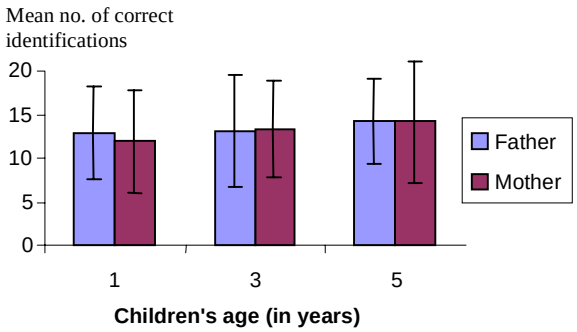

Figure 1. Mean number of correct identifications (out of 30) of parents of children at various ages (one-SD error bars). There is no significant difference in the level of correct identification of mothers versus fathers based on children's faces.

To reiterate, our analyses showed *no significant difference between the level of correct identification of mothers and the level of correct identification of fathers from children's faces* (Figure 1 and Table 1). Christenfeld and Hill (1995) did not perform a direct comparison between levels of identification of mothers and fathers. They simply compared the level of identification of mothers and fathers to the chance level of 33.3 percent by means of student t-tests, the items being the random factor. We also carried out this analysis for our data by comparing the mean number of identifications to chance (1/3 X 30 subjects = 10). At all ages tested, our results indicate that, while correct identification of mothers and fathers was significantly, although not overwhelmingly, higher than chance, there is no significant difference between the degree of father-identification and mother-identification.

It is particularly important to note that while the degree of correct association of parents with children is anywhere between 7 and 14% higher than chance, it remains surprisingly poor. In all cases, *non-identification exceeds 50%.*

Discussion

Present results do not replicate those of Christenfeld & Hill's (1995) study. Young children aged 1, 3 and 5 do not appear to resemble their fathers significantly more than they resemble their mothers.

It could be objected that the sample of faces used in this experiment is not a representative one. In fact, there is no clear reason why our sample of items would not be representative of the larger Caucasian population in general, and, crucially would be less representative than Christenfeld and Hill's original sample. Indeed, we used photographs from 28 families, whereas Christenfeld and Hill's stimuli were drawn from 24 families. Our stimuli were collected in the same way as those in the Christenfeld and Hill study, i.e. by asking friends, colleagues and acquaintances for photographs. We do not see any a priori reason why such a procedure would lead to the construction of an unrepresentative set of faces.

Age	Parent	Mean no. of identifications (SD in parentheses)	Student *t*	*p*
1	Father	12.893 (5.363)	2.854	<.01
	Mother	11.929 (5.937)	1.719	<.05
3	Father	13.178 (6.464)	2.602	<.01
	Mother	13.321 (5.644)	3.114	<.01
5	Father	14.143 (4.859)	4.512	<.001
	Mother	14.143 (6.996)	3.134	<.01

Table 1. Mean number of correct identifications (out of 30) as a function of the children's age and the parent's sex. Standard deviations are in parentheses. Note the absence of any significant difference in levels of correct identification of fathers and mothers based on a child's facial appearance.

Is our failure to replicate Christenfeld & Hill possibly attributable to an inappropriate sample of pictures that allowed *no* null hypothesis to be rejected? This would be very unlikely, because in *all six cases* of mothers and fathers for 1, 3, and 5 year old children, we found that the resemblance of parent to child is, as one would expect, significantly better than chance. In short, our sample *did* demonstrate a significant resemblance between parents and children, but *not that there was a significantly greater* resemblance between fathers and their children compared to mothers and their children. This means that our failure to find a significant difference in the resemblance of fathers-to-children versus mothers-to-children was not simply due to an insufficient amount of detail to be able to make resemblance assessments of any kind.

Experiment 2

Subjects

Forty-four undergraduate volunteers (22 females, 22 males) participated in the experiment.

Stimuli and materials

Thirty-two Caucasian families provided three photographs: one photograph of a child at one year, one photograph of the mother and one photograph of the father taken when the child was approximately one year old. For sixteen families, the child was a girl, for the other sixteen families the child was a boy. The stimuli presented to subjects were scanned versions of these photographs (size = 5.5 x 4.5 cm) of faces. None of the faces had glasses, beards or moustaches. Stimuli were presented using E-prime® on a PC.

Procedure

On each trial, participants were presented with the face of a child and, according to the condition, the faces of

three women or three men. All photographs were displayed on the computer screen. Their task was to identify as quickly and as accurately as possible the child's parent among the three presented adult faces. Participants responded by pressing a key on the numeric keypad of the computer keyboard (1 = left photo choice, 2 = middle photo choice and 3 = right photo choice). There were 32 trials (16 different girls and 16 different boys). The position of the real parent among the three adult photos was appropriately randomized. Each participant was presented with the 32 sets consisting one child and three possible parents in a different random order. The experiment was preceded by a short practice session using four trials that were not employed later in the experiment.

Results

The experiment had a repeated measures design with two factors : the gender of the parent and the gender of the child. One item was removed because the proportion of correct mother-infant identification was below the 2SD cutoff.

The first dependent measure was the proportion of correct identification of the parent. A 2 (gender of the parent) X 2 (gender of the child) ANOVA with repeated measures on both factors revealed no main effect of the gender of the parent ($F(1,43) < 1$), no main effect of the gender of the child ($F(1,43) < 1$), and no interaction effect ($F(1,43)<1$). See Table 2.

Parent	Infant gender	
	Girl	Boy
Mother	0.41 (0.16) 2.72 secs (0.85)	0.39 (0.16) 2.72 secs (0.88)
Father	0.38 (0.19) 2.82 secs (1.0)	0.38 (0.19) 2.73 secs (1.2)

Table 2. Mean proportions of correct identifications of parents, and mean correct RTs (in seconds) as a function of the gender of the child and the parent. Standard deviations are given in parentheses.

A control analysis taking the items as the random factor was also carried out and revealed the same pattern of results: none of the main and interaction effects were significant (all Fs <1).

The level of identification of mothers and fathers was also compared to the chance level of 33.3 percent by means of student t-tests, the subjects being the random factor. The overall mean level of correct identification of both mothers ($m = 0.397$; $t(43) = 3.696$; $p < .001$) and fathers ($m = 0.381$; $t(43) = 2.436$; $p < .05$) was significantly higher than chance.

The second dependent measure was the response latency (RTs). Mean correct recognition RTs to the mother and the father were computed for each subjects and submitted to a 2 (gender of the parent) X 2 (gender of the child) ANOVA. Five subjects were excluded from this analysis: two subjects whose RTs were particularly slow (RTs > 2 SD from the sample average) and three subjects who did not provide any correct recognition in one subcategory of items. This analysis revealed no main effect of the gender of the parent ($F(1,38)<1$, $p=0.49$), no main effect of the gender of the child ($F(1,38)<1$; $p=0.53$), and no interaction effect ($F(1,38)<1$; $p=0.71$). See Table 2.

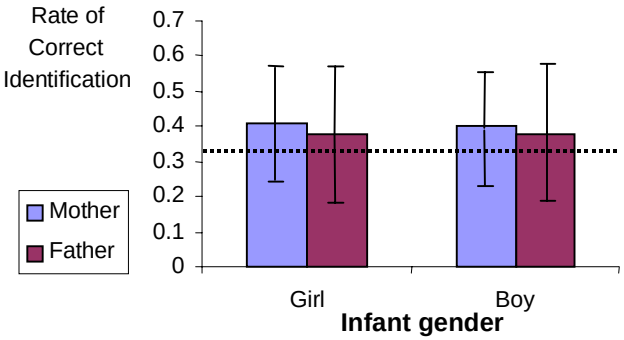

Figure 2. As in the first experiment, Exp. 2 shows no significant difference between child-father and child-mother rates of correct identification (1 SD error bars). The dotted line indicates chance level of identification.

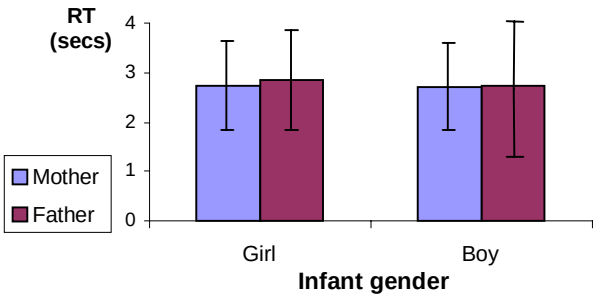

Figure 3. There is no significant difference between child-father and child-mother reaction times for correction identifications (1 SD error bars).

A control analysis taking the items as the random factor was also carried out and revealed the same pattern of results: none of the main and interaction effects were significant (all F's < 1).

Discussion

This second experiment was, above all, designed to repeat and refine the results obtained in Experiment 1. Entirely new stimuli (i.e., black-and-white photographs of adults and infants) were used. And, unlike the first experiment in which children of ages 1, 3 and 5 were used, here we focused exclusively on one-year old infants. (This was because the claim of Christenfeld and Hill bears specifically on one-year old infants: it is at that age that infants supposedly resemble their fathers more closely than their mothers.) The stimuli in this experiment were presented on a computer monitor instead of using the actual photographs, as in the first

experiment. And, of course, all of the participants were different from the first experiment.

All of the results of the first experiment were reproduced in this second experiment. As in the first experiment, we found that the level of correct identification of infant-father pairs was not significantly higher than that of infant-mother pairs. We also found that, as in the first experiment, that both the levels of infant-father and infant-mother identification are significantly above chance, as one might expect. Finally, we found no significant difference in the reaction times for correct responses for both infant-father and infant-mother pairs. In other words, there is no significant difference in the speed with which people can correctly identify an infant's mother or its father.

As in Experiment 1, these results are in clear contradiction with the results of Christenfeld and Hill (1995).

General Discussion

The evolutionary analysis of Christenfeld and Hill is based on the supposed advantage to one-year old babies of looking more like their father than their mother in order to encourage greater resource investment on the part of the father, thereby improving their chances of survival. This theory is certainly appealing, but we believe it is undermined by a number of considerations that we will review below.

We must begin by returning to the fundamental postulate of Darwinian evolution, namely that the ultimate winners in the game of evolutionary competition are those individuals who succeed in passing on the greatest amount of their genetic material to subsequent generations. Now, there would be little obvious evolutionary pressure for a child to resemble its mother, since the maternity of a child is never in doubt. This allows us to take the degree to which a child resembles its mother as a baseline of parent-child resemblance.

The essence of the argument against greater resemblance between fathers and their infants as opposed to infants and their mothers is based on the following simple observation: If father-child resemblance was strong enough to enable a father to be certain when a child was his, it would presumably also permit a father to identify that a child *was not his* (Brédart & French, 1999). Now, in the event that a child was not his, the chances of his withholding resources from the child (or very possibly killing the child outright) would be high. Even today, step-children are far more likely to be killed by step-parents than by natural parents. In the U.S. in 1976, for example, Daly and Wilson (1988) reported that children living with one or more substitute parents were *sixty-five times* as likely to be fatally abused as children living with their biological parents. Other studies report similar patterns of child mistreatment (for a recent short review see Daly and Wilson, 1996). Animal research has also clearly demonstrated the prevalence of infanticide by male rodents, carnivores and, in particular, primates (Hdry, 1979).

For much of the two-million year pre-agricultural course of human existence, three important conditions prevailed: male parental investment (Trivers, 1972) was necessary to ensure the survival of offspring, males were unable to completely control all possible sexual contact of their mates, and, finally, few individual males were able to provide resources for many females (Symons, 1987). Under these conditions, if babies had unambiguously resembled their fathers, a highly monogamous society would likely have emerged because few females would have risked the possibility of fathering another male's child, given that the bastard child would have been recognized as not belonging to her "official" (investing) mate (see also comments by R. Dawkins and other discussants following a paper by Wilson and Daly, 1997) and would thus have risked maltreatment and, quite possibly, death. In short, few females would have engaged in extra-pair copulation (EPC). However, in reality, this is contradicted by the fact that occasional EPCs by both sexes seems to be a universal feature of monogamous species (Mock and Fujioka, 1990), including humans. For example, rates of human misassigned paternity (based on blood typing tests) of 6-30% have been reported in studies done in southern England (Edwards, 1957; Philipp, 1973), 9% among the Venezuelan Yanomanö (Neel and Weiss, 1975; Smith, 1984), and 10% in rural Michigan, (Smith, 1984). Baker and Bellis (1995) have estimated a cross-cultural median EPC figure of 9%, with a range from 1.4-30%. Further, in a survey of 2078 English women, Bellis and Baker (1990) found that extra-pair copulations are significantly more likely to be timed just before ovulation than in-pair copulations. From his model of parent-infant resemblance, Pagel (1997) recently concluded that "even small amounts of paternity uncertainty are sufficient to select against parent-infant resemblance" (p.973).

Moreover, if relatively high father-child resemblance were the norm, evolution would tend to produce progressively greater degrees of father-child resemblance because any degree of resemblance significantly below that norm would engender suspicions on the part of the resource-providing male concerning the child's paternity. This would likely lead to a higher degree of resource-withholding than if the child had unambiguously resembled the father, which would ultimately translate into a lower rate of survival among those children who did not closely resemble their fathers. In other words, once evolution had established a trend of father-child resemblance in excess of baseline resemblance, there would be evolutionary pressure towards ever greater resemblance. One would therefore expect, after three million years of selection, that there would now be a

very *strong* tendency of father-child resemblance with respect to mother-child resemblance. However, our results — as well as those by Christenfeld and Hill — demonstrate that this is not the case. Indeed, in Christenfeld and Hill's data correct identification of fathers from infant faces occurred only in 49.2 percent of cases. In the two experiments reported in the present paper, the mean rate of correct identification for the father's of one-year-old children was only 10% higher than chance in the first experiment and 5% higher than chance in the second.

For these reasons, we believe that the original results reported by Christenfeld and Hill (1995) of greater father-child than mother-child resemblance in young children are most likely incorrect.

Conclusion

We believe that the experimental results presented by Christenfeld & Hill (1995) are most likely in error. We have attempted on three separate occasions to reproduce their results, each time with new photographic stimuli and new participants. We have used two separate measures (percentage of correct identifications and reaction times for correct identifications). In all cases, we have seen no evidence whatsoever of the results reported in their paper. In this paper we report two of our experiments. In addition, we provide a theoretical justification of the outcome of our experiments. We believe that the evidence presented in this paper casts serious doubt on the originally published study by Christenfeld and Hill.

Acknowledgments

The present paper was supported in part by Belgian Federal research grant IUAP P4/19 and research grant ARC 99/04-246 from the government of the French-speaking community of Belgium.

References

Baker, R. & Bellis, M. (1995). *Human Sperm Competition*. London: Chapman and Hall, 142-144.

Bellis, M. & Baker, R. (1990). Do Females Promote Sperm Competition? Data for Humans. *Animal Behaviour, 40*, 997-999.

Bennett, E. & Calvin, M. (1964). Failure to train Planarians reliably. *Neurosciences Research Program Bulletin*, July-August, 3-24.

Brédart, S. & French, R.M. (1999). Do babies resemble their fathers more than their mothers? A failure to replicate Christenfeld & Hill (1995). *Evolution and Human Behavior, 20*(3), 129-135.

Byrne et al., (1966). Memory transfer. *Science, 153*, 658-9.

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

Daly, M. & Wilson, M. (1988). *Homicide*. Hawthorne, NY: Aldine de Gruyter.

Daly, M. & Wilson, M. (1996). Violence against stepchildren. *Current Directions in Psychological Science, 5*, 77-81.

Edwards, J. H. (1957). A critical examination of the reputed primary influence of ABO phenotype on fertility and sex ratio. *British Journal of Preventive and Social Medicine, 11*, 79-89.

Gaulin, S. & Schegel, A. (1980). Paternal confidence and paternal investment: a cross-cultural test of a sociobiological hypothesis. *Ethology and Sociobiology, 1*, 301-309.

Hdry, S. (1979). Infanticide among Animals: A Review, Classification and Examination of the Implications for the Reproductive Strategies of Females. *Ethology and Sociobiology, 1*, 13-40.

Hilgard, E., Atkinson, R. & Atkinson, R. (1971). *Intro. to Psychology*. Harcourt, Brace Jovan. 232-233.

McConnell, J. (1962). Memory transfer through cannibalism in Planarians. *J. Neurophys, 3*, 42-8.

Mock, D. & Fujioka, M. (1990). Monogamy and long-term pair bonding in vertebrates. *Trends in Ecology and Evolution, 5*, 39-43.

Munn, N., Fernald, L. & Fernald, P. (1969). *Intro. to Psychology*. Houghton Mifflin. 268-269.

Neel, J.V. & Weiss, M. (1975). The genetic structure of a tribal population, the Yanomama Indians. XIII. Biodemographic studies. *American Journal of Physical Anthropology, 42*, 25-51.

Pagel, M. (1997). Desperately concealing fathers: a theory of parent-infant resemblance. *Animal Behaviour, 53*: 973-981.

Philipp, E. (1973). Discussion: moral, social and ethical issues. In *Law and Ethics of A.I.D. and Embryo Transfer. Ciba Foundation Symposium* (Vol. 17), G.E.W. Wostenholme and D.W. Fitzsimons (Eds.). Amsterdam: Elsevier, Excerpta Medica, North-Holland, 63-66.

Smith, R. L. (1984). Human sperm competition. In *Sperm Competition and the Evolution of Animal Mating Systems*, R.L. Smith (Ed.). London: Academic Press, 601-660.

Symons, D. (1987). An evolutionary approach: can Darwin's view of life shed light on human sexuality? In *Theories of Human Sexuality*, J. H. Geer and W. O'Donohue (Eds.) NY: Plenum, 91-125.

Trivers, R. (1972). Parental Investment and Sexual Selection. In *Sexual Selection and the Descent of Man*, B. Campbell (Ed.). Chicago: Aldine de Gruyter, 136-179.

Wilson, M. & Daly, M. (1997). Relationship-specific social psychological adaptations. *In Ciba Foundation Symposium 208: Characterizing Human Psychological Adaptations*, G.R. Bock and G. Cardew (Eds.). Chichester, UK: John Wiley & Sons, 253-268.