Heterosis Doesn’t Cause the Flynn Effect: A Critical Examination of Mingroni (2007)

Michael A. Woodley
Royal Holloway, University of London

Mingroni (2007) proposed that heterosis or hybrid vigor may be the principal driver of the Flynn effect—the tendency for IQ scores to increase at a rate of approximately 3 points per decade. This model was presented as a resolution to the IQ paradox—the observation that IQ scores have been increasing despite their high adult heritability—on the basis that substantial changes in IQ can only be accounted for by changes in underlying genetic factors. It is here argued that this model is predicated upon a misconception of the Flynn effect, which is most pronounced on the least g-loaded components of cognitive ability tests and is uncorrelated with genetic effects such as inbreeding depression scores (which are correlated with the g loadings of tests). Evidence supportive of the recently proposed life history model of the Flynn effect is presented. In the discussion, other theoretical objections to the heterosis model are also considered. On this basis, it is concluded that the Flynn effect is strongly entwined with developmental status and that heterosis cannot be its principal cause.

Keywords: heterosis, Flynn effect, IQ paradox, life history, Mingroni

Mingroni (2007; see also Mingroni, 2004) presented an intriguing model concerning the causes of the mysterious Flynn effect—namely, the tendency for IQ scores to increase at a rate of about 3 points per decade, which has been observed in a number of mostly industrialized countries (Flynn, 1984, 1987). The model is predicated upon the idea that heterosis, or hybrid vigor, might be a major driver of the effect. Heterosis is described as “a genetic effect that results from matings between members of genetically distinct subpopulations, such as has been occurring in human populations through the breakup of small, relatively isolated communities owing to urbanization and greater population mobility” (Mingroni, 2007, p. 806) and is the opposite of inbreeding depression which is associated with the combination of rare deleterious recessive alleles that reduce fitness (Bittles, 2010). Inbreeding at the level of first cousins is known to depress IQ scores by between 2.5 and 3.5 points per generation on average (Jensen, 1983), and there is evidence that outbreeding can in some instances raise IQ (Nagoshi & Johnson, 1987). Furthermore, IQ is known to have a high adult heritability (approximately .75; Neisser et al., 1996), despite which massive gains have been occurring amongst populations—a phenomenon referred to as the IQ paradox (Dickens & Flynn, 2001). Mingroni argued that given the high adult heritability of IQ, substantial changes in the trait would require genetic changes within populations, which can result from heterosis—therefore, heterosis presents a potential resolution to this paradox. Mingroni is not the first to advance the hypothesis that heterosis may contribute to the Flynn effect (see Jensen, 1998); the novelty of his argument stems from the large role that he believes heterosis may play in facilitating the effect.

Here it is demonstrated that Mingroni’s model is theoretically flawed. This is principally because the secular gains associated with the Flynn effect do not concern g on tests of cognitive ability. Heterosis is a purely genetic effect, and it has been observed that genetic effects such as inbreeding depression scores on tests of cognitive ability correlate significantly with their g loadings and not at all with secular gains. Thus, as the Flynn effect concerns neither g nor genetic effects, it is concluded that the arguments bought to bear by Mingroni in defense of his hypothesis cannot be considered relevant to the question of what causes the Flynn effect—a question that is still very much open to debate.

Dissecting the Flynn Effect

The Flynn effect manifests itself predominantly on fluid intelligence subfactors (i.e., nonverbal abilities such as educing patterns; Colom, Andres-Pueyo, & Juan-Espinosa, 1998; Emanuelsson, Reuterberg, & Svensson, 1993; Emanuelsson & Svensson, 1990; Flynn, 1987, 1998, 2006; Jensen, 1991; Loehlin, 1996; Lynn, 2009; Lynn & Hampson, 1986, 1989; Teasdale & Owen, 2000); however, there is evidence from a handful of studies of gains in crystallized intelligence subfactors also (such as verbal ability; e.g., Pietschnig, Voracek, & Formann, 2010; Römlund & Nilsson, 2009; Utl & Van Alstine, 2003; Voracek, 2006; Wicherts et al., 2004).

The Flynn effect therefore concerns differential gains in abilities rather than uniform gains across abilities, which is what one would
expect if the effect were associated with gains in the latent factor $g$. Although this position has been subject to debate (e.g., Blair, 2006; Colom, Juan-Espinosa, & García, 2001; Flynn, 1999a, 1999b, 2000; Jensen, 1996, 1998; Must, Must, & Raudik, 2003; Rushton, 1999, 2000; Voracek, 2006), a handful of studies in need of special mention have done much to help clarify the issues. Wicherts and coworkers (2004), employing multigroup confirmatory factor analysis, found that the assumption of factorial invariance with respect to cohorts, which is a necessary criterion for the invocation of changes in the level of $g$ in the Flynn effect, is in fact untenable. They found that this was true with respect to five data sets for which IQ data were available on two cohorts. Mingrioni (2007), in discussing this finding, conceded that “this finding would appear to be inconsistent with heterosis as the sole cause of the rise in IQ” (p. 812); however, he suggested that “the failure to observe complete measurement invariance between cohorts does not allow one to preclude the possibility that heterosis is a partial or even major cause of the Flynn effect” (Mingrioni, 2007, p. 812).

It is acknowledged that this is theoretically reasonable; however, a more recent study by Rushton and Jensen (2010) casts doubt on the validity of even a partial contribution to the Flynn effect from heterosis. They found that $g$ loadings and genetic effects such as inbreeding depression scores on the 11 subtests of the Wechsler Intelligence Scale for Children correlated either significantly negatively or not at all with secular gains due to the effect ($r = −.33$, $p < .001$, and .13, ns, respectively). This indicates that there is no connection between secular gains and genetic effects and that the effect is not occurring on the most $g$-loaded components of the tests. The finding that genetic effects are positively associated with $g$ loadedness and not at all with secular gains has been demonstrated in previous work by Rushton (1999), who found that whereas secular gains on the Wechsler Intelligence Scale for Children—Revised and Wechsler Intelligence Scale for Children—Third Edition cluster (in principal component analysis), this is independent of the cluster associated with genetic effects such as inbreeding depression scores or effects thought to contain a genetic component such as Black–White differences and $g$-factor loadings, suggesting that while secular gains are robust, there is no genetic basis to the Flynn effect (see also Rushton, 2000). This is completely contrary to what one would expect if heterosis were “a partial or even major cause of the Flynn effect” (Mingrioni, 2007, p. 821).

There are other findings that seem to disconfirm the heterosis hypothesis also. For example, the study of Sundet, Eriksen, Børren, and Tambs (2010), which investigated the Flynn effect with respect to the role of age differences between siblings in a large Norwegian sample of brother pairs, found that the effect can be predicted from the effects of age differences between brothers on their IQ scores, suggesting a significant role for factors operating within families. This significantly weakens the case for it resulting from factors operating solely between families such as heterosis, at least within Norway. In another study, Woodley (2009) investigated whether cross-national variability in per capita levels of consanguineous marriage (marriage involving individuals related to one another at the level of second cousin or closer) was in anyway predictive of cross-national variability in the IQ scores reported by Lynn and Vanhanen (2006). Although a significant positive correlation was found between the two variables ($r = −.62$, $p < .01$), this correlation was not robust to the effects of controlling for gross domestic product (GDP) per capita, levels of democracy, and education in multiple regression. In point of fact, consanguinity turned out to be wholly independent of IQ ($β = 0$, $p > .05$). Given that a compelling case has been made for the cross-national IQ estimates of Lynn and Vanhanen being heavily conflated with potential causes of the Flynn effect (Wicherts, Borsboom, & Dolan, 2010; Wicherts, Dolan, Carlson, & van der Maas, 2010; Wicherts & Wilhelm, 2007), if heterosis were substantially involved in the effect, one might expect a measure of inbreeding to be a somewhat robust predictor of cross-national differences in IQ; however, this is simply not the case.

It finally needs to be noted that Mingrioni himself also appears to have found evidence directly disconfirming the heterosis hypothesis, which was presented at the 2002 International Society for Intelligence Research conference (see Mingrioni, 2002). In this study, he examined the effects of intergenerational differences in genetic heterozygosity on IQ amongst a sample of 92 parent–child trios. The heterosis hypothesis requires that the higher IQ amongst the children should be positively associated with higher levels of heterozygosity; however, the opposite was found, namely, that the parents exhibited higher levels of heterozygosity than the children despite the children exhibiting higher levels of IQ, as measured by SAT scores (Mingrioni, 2002). An implication of this finding is that secular gains on IQ tests are wholly independent of intergenerational differences in the degree of heterozygosity, thus indicating that the Flynn effect cannot be genetic in origin.

### Life History as a Potential Alternative Causal Model

What factors may plausibly be involved in facilitating the Flynn effect? It would appear that a number of environmental and developmental factors are likely involved. Studies have found evidence for contributions from nutrition (Colom, Llusí-Font, & Andres-Pueyo, 2005; Lynn, 2009), family structure, pathogen stress/health and education (Daley, Whaley, Sigman, Espinosa, & Neumann, 2003; Eppig, Fincher, & Thornhill, 2010), and cognitive stimulation (Colom, Flores-Mendoza, & Abad, 2007). This is by no means a complete review of the diverse literature on this topic. A recently proposed model invokes life history as a potential causative factor (Woodley, in press-b). Life history speed describes individual differences in preferences toward the development of either a fast life history strategy characterized by high mating effort (resources allocated toward mating) or a slow life history characterized by high parental effort (resources allocated toward offspring survival) and high somatic effort (resources allocated toward growth and maintenance; Figueredo, Vásquez, Brumbach, & Schneider, 2004). This model is based on the cognitive differentiation–integration effort hypothesis (Woodley, in press-a), which holds that individual differences in life history speed are associated with a tradeoff between the strength of the positive manifold ($g$) and the development of separate abilities. Integrated abilities benefit individuals coping with unstable environments and people (such as the short-term mating market) as they confer ecological flexibility via a heightened capacity to cope with situational demands, whereas differentiated abilities permit socioecological specialization in response to competition. This tradeoff occurs independently of the level of $g$, which has been found not to correlate with measures of the latent life history factor ($K$) at individual differences scales (Woodley, in press-a). Integra-
tion effort and differentiation effort therefore constitute specialized forms of mating effort and somatic effort (with a possible contribution from parental effort also), respectively. As the Flynn effect appears to be associated with secular declines in the strength of $g$ coupled with secular gains in abilities (Juan-Espinosa, Cuevas, Escorial, & García, 2006; Kane, 2000; Kane & Oakland, 2000; Lynn & Cooper, 1993, 1994), the effect may result from differentiation effort investment in response to factors that have historically slowed the life history speed of populations, such as the move toward smaller families, better nutrition, generalized education, and disease prevention, by lowering infant mortality, thus providing a framework within which the effects of seemingly disparate developmental influences can be integrated. As mentioned in the previous section, national IQs are conflated with developmental indicators believed to be associated with the Flynn effect; therefore, they likely substantially capture differences in the rate at which secular gains have been occurring between nations and do not wholly reflect national differences in $g$ (Wicherts, Borsboom, & Dolan, 2010; Wicherts, Dolan, et al., 2010; Wicherts & Wilhelmi, 2007; cf. Rindermann, 2007). Given this, measures of educational, infectious disease, family size, and nutrition should all load along with national IQ on the cross-national life history $K$ superfactor identified by Templer (2008). To test this, data on birth rate (a family size proxy), life expectancy, infant mortality, GDP, and national IQ (from Templer, 2008), along with education index (United Nations Development Program, 2010) and log-transformed disability-adjusted-life-years infectious diseases and nutrition (World Health Organization, 2004), were collected for 127 countries. Principal component analysis (using SPSS) revealed the presence of a single $K$ superfactor accounting for 82% of the variance. This finding therefore provides support for the hypothesis that life history may constitute a factor common to the influence of diverse developmental variables on the Flynn effect.

**Discussion**

The evidence presented here strongly suggests that heterosis is in no way associated with the Flynn effect as has been typically measured and described in the literature. Even if heterosis did play a small role, its contribution to raising IQ would have been heavily attenuated by much bigger secular gains in abilities bought about by the influence of diverse developmental factors. In Mingroni’s (2007) Parma Valley simulations, in which he attempted to estimate IQ gains over the course of 50 years by assuming heightened demographic mobility, he noted that “all five simulations predict an increase in IQ of about 3 points (0.2 standard deviations)” (p. 822). As Lynn (2009) and Flynn (2009) both pointed out, however, gains of 2 or 3 points due to heterosis are a far cry from the 27-point increase that has been observed in the United States over the last 80 or so years. The only mechanism through which heterosis might substantively influence the Flynn effect might be through a social multiplier effect involving small genetic gains in $g$ having a larger positive effect on developmental factors that might encourage further secular gains amongst abilities (Woodley, 2009). A significant impediment to this model lies in the observation that recent significant inbreeding does not appear to be a characteristic of many of the nations in which the Flynn effect has occurred, however (Flynn, 2009).

Another theoretical objection to the model concerns the idea that heterosis would actually lead to increases in IQ in all instances. Although it is likely that it would in instances where inbreeding is a relatively novel constraint on populations, there is no reason why it should in populations where inbreeding has been practiced constantly for many generations and where there has been an opportunity for purifying selection to purge the worst allele combinations. This mechanism is known in the nonhuman animal literature; for example, within the inbred population of the Scandinavian wolf, purifying selection against deleterious alleles has been observed with respect to a number of genetic markers (Hagenblad, Olsson, Parker, Ostrander, & Ellegren, 2009). Inbreeding is likely evolutionarily basic to humans and is still practiced in many different cultures, with around 10% of all marriages globally involving first cousins (Bittles, 2010). One evolutionarily significant factor possibly involved in shaping this reproductive behavior may be pathogen stress, which requires that locally coadapted immunogene complexes be maintained at a high frequency within populations as part of a coevolutionary arms race with locally occurring parasites (Demic & Nicholls, 2007; Fincher & Thornhill, 2008; Hoben, Buunk, Fincher, Thornhill, & Schaller, 2011).

Another point in need of addressing concerns the IQ paradox. This is evidently a red herring as it would only be a paradox if the gains from the Flynn effect were due to increases in $g$, which has high adult heritability. As the effect only concerns the non-$g$ variance unique to specific cognitive abilities, environmental factors can be ruled back in as potential causes, as these have substantively lower heritabilities than $g$ (Carroll, 1993) so are therefore much more amenable to environmental manipulation. There is in fact evidence that IQ has been in decline in industrialized nations as a consequence of the existence of negative correlations between IQ and fertility (Lynn, 1998) and that this effect is associated with $g$ (Meisenberg, 2010). Again, this is hard to reconcile with the heterosis model as how can $g$ be simultane-

---

1 Templer (2008) included a measure of skin reflectance in his original $K$ superfactor. Attempts to determine the historical robustness of the factor via reconstruction using 1997 data for 41 countries found that this variable loaded preferentially on a second principal component. On this basis, it was concluded that skin reflectance is not robustly associated with changes in life history speed indicators, hence it was not included here.

2 Principal axis factoring produced a similar result. The Kaiser-Meyer-Olkin measure of sampling adequacy was .992, indicating that the correlations were adequate for factor analysis; additionally, Bartlett’s test of sphericity falsified the null hypothesis (i.e., no correlations between the variables).
References


HETEROSIS DOESN’T CAUSE THE FLYNN EFFECT


Must, O., Must, A., & Raudik, V. (2003). The secular rise in IQs: In Estonia, the Flynn effect is not a Jensen effect. *Intelligence, 31*, 461–471.


Received January 31, 2011
Revision received May 18, 2011
Accepted May 27, 2011