

# Rethinking the Flynn Effect

## 1. Introduction

During the twentieth century, several researchers noticed that overall raw performance on intelligence exams seemed to be generally increasing over time. Then it was James Flynn in the 1980s who provided abundant evidence that this phenomenon was essentially universal, thereby drawing greater attention to it, and the phenomenon would be eventually dubbed the Flynn effect (Flynn, 1984, 1987). Since its discovery, the Flynn effect has remained a scientific mystery, an unexpected result defying the many attempts to explain it (Trahan et al., 2014). And perhaps due in part to this intransigence, the Flynn effect has also been the target of an unease and mistreatment, everything from an initial dismissiveness of the phenomenon's significance by James Flynn himself, to an ongoing tendency towards overfitting by models distinguished for their parameterized complexity, to a more recent groundswell of anticipation for the Flynn effect's apparently imminent demise.

Ironically, the one approach to the Flynn effect that seems not to have been given serious consideration by the intelligence research community is to embrace the phenomenon as a fundamental and enduring property of human intelligence, as consequential in impact as say Spearman's *g* (general intelligence ability). I believe this neglect will prove to be a mistake. Afforded a reasonable degree of presumptive acceptance, the Flynn effect emerges as not a twentieth-century aberration, not a temporary quirk to be solved with labored explanations. Instead, the Flynn effect can be demonstrated as having been with humanity for an extremely long period of time, ever since the beginning of the human behavioral transformation, and there is no reason to expect that the Flynn effect will end anytime soon. Thus, far from being a candidate for unease and mistreatment, the Flynn effect should be recognized as foundational to a complete understanding of human intelligence.

## 2. Various Forms of Dismissal

In reading through the intelligence research literature, one seldom comes across the Flynn effect being described with such terms as *foundational*, *fundamental* or *permanent*. Instead, the words *aberration*, *not real* and, above all else, *temporary* are more frequently applied. One of the earliest attacks on the Flynn effect came from James Flynn himself, who having inferred that the IQ gains he was seeing in the data would classify his nineteenth-century ancestors as mentally deficient and would render today's children as observably smarter than their parents, opined that the Flynn effect was not indicative of a true increase in human intelligence (Flynn, 1987, 1998, 2007).

Flynn would later waver in this opinion, eventually devoting many articles and entire books to exploring the Flynn effect as a serious topic for human intelligence (Flynn 2007, 2012), even going so far as to collaborate with William Dickens to develop a labyrinthine model for explaining how the Flynn effect can be reconciled to *g* (more on this in a moment). Nonetheless, throughout his career, Flynn never seemed to come to peace with his namesake subject, remaining confused by its apparent paradoxes and troubled by its disruptive implications. In this, Flynn was certainly not alone. Unease with the Flynn effect is practically palpable within the literature, almost the entirety of which can be summed up in the following manner:

1. The Flynn effect is puzzling and unexpected, to the point of being surely an aberration.
2. The Flynn effect began likely sometime during or just before the twentieth century, and certainly no earlier than the Scientific and Industrial Revolutions.
3. The Flynn effect cannot go on forever.
4. Therefore, there must be a cause (or a set of causes) that both explains how the Flynn effect recently came into being, and how the Flynn effect will of necessity soon go away.

The list of suggested causes is nearly endless—heterosis, better nutrition, expanded education—to name just a few (Mingroni, 2007; Lynn, 1989; Baker et al., 2015). But the problem has always been that these suggestions lack sufficient spatial and temporal reach to match the nearly ubiquitous impact of the Flynn effect. Thus, unable to explain the phenomenon easily, intelligence researchers have turned to two alternative approaches. The first approach has been simply to wait out the Flynn effect, to begin looking for the signs of its predicted and requisite end. And perhaps not coincidentally, no sooner has this strategy been adopted than the evidence has begun to pour in supporting the anticipated result, even to the point of considering online surveys as evidence for the Flynn effect's reversal (Pietschnig & Gittler, 2015; Dutton, van der Linden & Lynn, 2016; Dworak, Revelle & Condon, 2023).

The second approach has been to subdue the Flynn effect with complexity and multiplicity. There have been three prominent attempts along these lines. The first has been the multiple causation theory (Jensen, 1998), the suggestion that although no one cause by itself can produce the Flynn effect, many causes in combination can adequately do the trick; in essence, *everything* causes the Flynn effect. The second attempt has been the Dickens-Flynn model (Dickens & Flynn, 2001), a complex gathering of mathematical formulae and parameterized concepts—social multipliers, rolling triggers, amplified feedback loops—all tunable to whatever IQ data one might happen to obtain. Indeed, Dickens and Flynn have been known to tout their model's ample knobs and levers as one of its primary virtues (Dickens & Flynn, 2002). The third attempt has been the life history model of intelligence (Woodley, 2012), a conceptually intricate collection of statistically representable life history speed factors—such as pathogen stress, nutrition, family size, education, etc. These life history speed factors could presumably, and in the right combination, produce the Flynn effect, and then in a different combination make it disappear.

The problem with this second approach is the concept known as overfitting. It is well understood within the practice of data science that given enough complexity and/or enough free parameters, a model can always be developed that will fit snugly to any given set of data (Hawkins, 2004). Unfortunately, such models thereby lose all their explanatory and predictive power, rendering them empty of consequence. The multiple causation theory, the Dickens-Flynn model, and the life history model—these are all classic examples of overfitting. They tell us everything we could possibly want to know about the data, but leave it problematic as to whether they are telling us anything insightful about human intelligence.

### **3. Human History and the Flynn Effect**

The notion that the Flynn effect is only a recent phenomenon is contradicted by the entire course of human history. Although we are not exactly certain when the human species first began its turn towards behavioral modernity, any reasonable estimate would put this moment at no earlier than a few hundred thousand years ago (Henshilwood & Marean, 2003), and until that time, the human species was pure animal. Not a single concept to be found on a modern intelligence exam would have been present in the species—no language, no arithmetic, no

geometric patterns, no logic, no anything. Try to imagine administering Wechsler to one of those ancient humans—it would be no more successful than administering Wechsler to a wild leopard. But this means that measurable human intelligence was at that time quantifiable as absolute zero, and since measurable human intelligence today has clearly advanced to a more substantive number, that overall increase, *by definition*, is a Flynn effect. Indeed, it is a *massive* Flynn effect.

Furthermore, the increase in measurable intelligence over that time has been continuous and not sudden. For instance, by the time of the out-of-Africa hunter-gatherers, the human species had begun to display an observable degree of intelligence—controlled fire, crafted weapons, ornamental jewelry, cave paintings, etc. (Klein, 2002). Administering an intelligence exam to that population would be conceivable, although the contents of such an exam would have to be simple and crude by modern standards (because of limited vocabulary, primitive numeracy, etc.). Sometime later, by the era of the early farmers of the Fertile Crescent, a more advanced level of intelligence had become apparent within the species—permanent abodes, irrigation techniques, pottery making (Christian, 2018)—and an intelligence exam appropriate for that population would be more sophisticated and more varied than the one appropriate for the hunter-gatherer population; and yet still, such an exam would have to remain simple by modern standards, because writing and arithmetic, for instance, were still on the verge of being invented. It has only been through recent times, an era increasingly dominated by such artifacts as books, cities and automobiles, that humans have gained enough proficiency to allow them to tackle the complexities of Stanford-Binet and Wechsler. Thus, the current level of measurable human intelligence has not come into existence suddenly, but instead has been steadily progressing ever since the first days of the human turn, and this means that the Flynn effect has been operative within the human population for a very long time.

And although conceptualizing intelligence exams for ancient populations does require some imaginative reasoning, do note that this reasoning remains entirely consistent with what actually took place during the twentieth century, on actual intelligence exams. The only way that an average IQ test taker from the late twentieth century could have equaled the raw score of an average IQ test taker from the early twentieth century would have been for the later exams to be altered to be more varied, more complex and more difficult (which in many instances they actually were). Increasing sophistication in the contents of intelligence exams is a palpable indicator of a progressive increase in the overall level of measurable intelligence, an increase that was evident in actuality during the twentieth century, and has been demonstrably evident across the entire course of human history, ever since the beginning of the human behavioral transformation.

But if this is indeed so, then it also puts to rest any notion that the Flynn effect is ending or reversing. What an incredible coincidence it would be to come across a phenomenon operative within the species for tens of thousands of years, and then now suddenly, right at the very moment of its conscious discovery, the phenomenon screeches to an abrupt halt—that makes no sense at all. Whatever has been driving the Flynn effect over the course of human history, including right through the entirety of the twentieth century, it must still be operative within the human population today.

#### **4. A Model of Acceptance**

In other works (Griswold, 2017, 2023), I have outlined a model of human intelligence that accepts the Flynn effect as valid, fundamental and enduring. The model's salient features include the following:

1. Measurable human intelligence is best described as the orthogonal product of two different factors: one, individual general intelligence ability (such as that quantifiable by *g*); and two, the total amount of artificial construction contained within the human environment, the target towards which general intelligence ability is applied.
2. The contents of an IQ exam are themselves artificial constructions— words, digits, sequences, puzzles, matrices, etc. As such, they serve as a proxy for the artificial construction contained within the human environment. That is, an individual's performance on an IQ exam is an indirect assessment of that individual's ability to navigate and to master the artificial construction to be found in that individual's everyday world.
3. The overall level of general intelligence ability remains stable within the population over time. This is exactly as to be expected for an ability driven primarily by genetic and neural characteristics.
4. On the other hand, the amount of artificial construction contained within the human environment has been continuously increasing ever since the beginning of the human behavioral transformation. Humans have progressed steadily (and successfully) from living in an entirely natural setting to living in a world now dominated by accruing amounts of artificial construction.
5. Since measurable human intelligence is the orthogonal product of these two different factors, one of which has been continuously increasing over time (and the other of which has remained entirely stable), measurable human intelligence has also been continuously increasing over time. This is a precise description of the Flynn effect, and it marks the accruing amount of artificial construction contained within the human environment as the sole driver and the sole explanation of the Flynn effect.
6. The above statements demonstrate that there is nothing contradictory (or paradoxical) about a stable general intelligence ability coexisting with increasing levels of measurable intelligence. Our nineteenth-century ancestors were not mentally deficient, nor are today's children smarter than their parents, because general intelligence ability remains stable over time. Nonetheless, later generations, by virtue of living in environments containing increased levels of artificial construction, and by applying their general intelligence ability to that increased level of artificial construction, will thereby demonstrate greater levels of performance on intelligence exams.
7. The above statements also demonstrate that the widespread presumption that intelligence is primarily a function of the human brain is in essence incorrect. The brain plays a role—by being responsive to artificial construction—but the locus of intelligence is the expanding structure of the artificial environment, and not the neurons of the human brain.
8. Since the contents of IQ exams serve as a proxy for the artificial construction contained within the human environment, and since the amount and type of that artificial construction is constantly changing over time, the contents of IQ exams must themselves be adjusted on a regular basis, typically towards greater variety, greater complexity and greater difficulty, to reflect the growing challenge humans must face in navigating the increasing amounts of artificial construction contained within their surrounding world.
9. Barring a catastrophe (such as civilization collapse), the amount of artificial construction within the human environment will continue to increase into the foreseeable future, and future generations will be obliged to navigate and to master this increasing amount of artificial construction, and will thereby also go on to demonstrate increasing levels of performance on future intelligence exams. Therefore, there is no reason to expect that the Flynn effect will end anytime soon.

If I were to characterize this model succinctly, I would say it is an effort to employ the Flynn effect to challenge our preconceptions about human intelligence, as opposed to an application of our preconceptions about human intelligence to wrestle the Flynn effect into submission.

## 5. Conclusion

From the history of science, we have an analogous circumstance to that of the current situation regarding the Flynn effect. In the late nineteenth century, the Michelson-Morley experiment was conducted in order to detect the presence of the luminiferous aether, by measuring the difference in the speed of light in the direction of Earth's motion versus the speed of light at right angles to that motion (Michelson & Morley, 1887). But the result of the experiment turned out to be entirely unexpected, with the speed of light measuring the same in every direction observed. Scientists spent the next twenty years flailing against this result, first with insistence that the experiment was flawed, and then with later implication that the verified outcome was incomprehensible or inconsequential or both (Swenson, 1970). On a different front, by around the year 1905, Hendrik Lorentz and Henri Poincaré had worked out the complex mathematics needed to reconcile a stationary aether to the presumed spatial and temporal contractions experienced by moving bodies, equations and models derived specifically to fit to the Michelson-Morley results (Lorentz, 1904; Poincaré, 1900). But as with all cases of overfitting, what the Lorentz/Poincaré equations and models revealed was only information about the data itself, not providing any useful elucidation about the processes *underlying* that data.

So how was this circumstance resolved? It was resolved when a naive young gentleman dared an approach that had eluded the scientists of that day, namely accepting the Michelson-Morley outcome as both valid and fundamental, postulating that the speed of light was indeed the same in every inertial frame, and then working out the consequences from there. I would encourage everyone to read Einstein's original paper on special relativity—it is a paradigm of simplicity and straightforwardness (Einstein, 1905). And because Einstein's approach was so simple and straightforward, it retained its predictive and explanatory power, unveiling a host of compelling insights into the characteristics of space, time, matter and energy.

I believe a similar fate awaits the Flynn effect. When researchers come to accept the phenomenon as valid, fundamental and enduring (which perhaps requires a certain amount of naiveté), they will have the context for a model of human intelligence not encumbered by too much parametric complexity. And as with all simple and straightforward models, this one will retain some of its predictive and explanatory force, perhaps unveiling useful insights into the course of human history and into the nature of human intelligence.

## References

Baker, D. P., Eslinger, P. J., Benavides, M., Peters, E., Dieckmann, N. F., & Leon, J. (2015). The cognitive impact of the education revolution: A possible cause of the Flynn Effect on population IQ. *Intelligence*, 49, 144–158. <https://doi.org/10.1016/j.intell.2015.01.003>

Christian, D. (2018). *Origin story: a big history of everything*. First edition. New York, Little, Brown and Company.

- Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: the IQ paradox resolved. *Psychological review*, 108(2), 346–369. <https://doi.org/10.1037/0033-295x.108.2.346>
- Dickens, W. T., & Flynn, J.R. (2002). The IQ paradox: still resolved. *Psychological Review*, 109, 764-71.
- Dutton, E., van der Linden, D., & Lynn, R. (2016). The negative Flynn Effect: A systematic literature review. *Intelligence*, 59, 163–169. <https://doi.org/10.1016/j.intell.2016.10.002>
- Dworak, E. M., Revelle, W., & Condon, D. M. (2023). Looking for Flynn effects in a recent online US adult sample: Examining shifts within the SAPA Project. *Intelligence*, 98. <https://doi.org/10.1016/j.intell.2023.101734>
- Einstein, A. (1905). On the Electrodynamics of Moving Bodies. *Annalen der Physik*, 17, 891-921.
- Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95(1), 29–51. <https://doi.org/10.1037/0033-2909.95.1.29>
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101(2), 171–191. <https://doi.org/10.1037/0033-2909.101.2.171>
- Flynn, J. R. (1998). IQ gains over time: Toward finding the causes. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 25–66). American Psychological Association. <https://doi.org/10.1037/10270-001>
- Flynn, J. R. (2007). *What is intelligence?: Beyond the Flynn effect*. Cambridge University Press.
- Flynn, J. R. (2012). *Are we getting smarter? Rising IQ in the twenty-first century*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139235679>
- Griswold, A. (2017). *Concerto for Intelligence*. iUniverse. (accessible online: <https://www.grizzalan.com/concertoformintelligence>)
- Griswold, A. (2023). *Autistic Rhapsody*. iUniverse. (accessible online: <https://www.grizzalan.com/autisticrhapsody>)
- Hawkins D. M. (2004). The problem of overfitting. *Journal of chemical information and computer sciences*, 44(1), 1–12. <https://doi.org/10.1021/ci0342472>
- Henshilwood, C. S., & Marean, C. W. (2003). The Origin of Modern Human Behavior: Critique of the Models and Their Test Implications. *Current Anthropology*, 44(5), 627–651. <https://doi.org/10.1086/377665>
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Praeger Publishers/Greenwood Publishing Group.
- Klein, R. (2002). *The Dawn of Human Culture*. New York: Wiley.
- Lorentz, H. A. (1904). Electromagnetic phenomena in a system moving with any velocity less than that of light. *Proc. Acad. Science Amsterdam*, IV, 669–78.

Lynn, R. (1989). A nutrition theory of the secular increases in intelligence; positive correlations between height, head size and IQ. *British Journal of Educational Psychology*, 59(3), 372–377. <https://doi.org/10.1111/j.2044-8279.1989.tb03112.x>

Michelson, A. A., & Morley E. M. (1887). On the relative motion of the Earth and the luminiferous ether. *American Journal of Science*, 34(203), 333–345.

Mingroni, M. A. (2007). Resolving the IQ paradox: Heterosis as a cause of the Flynn effect and other trends. *Psychological Review*, 114(3), 806–829. <https://doi.org/10.1037/0033-295X.114.3.806>

Pietschnig, J., & Gittler, G. (2015). A reversal of the Flynn effect for spatial perception in German-speaking countries: Evidence from a cross-temporal IRT-based meta-analysis (1977–2014). *Intelligence*, 53, 145–153. <https://doi.org/10.1016/j.intell.2015.10.004>

Poincaré, H. (1900). The theory of Lorentz and the principle of reaction. *Archives neerlandaises*, V, 252–78.

Swenson, L. S. (1970). The Michelson-Morley-Miller Experiments before and after 1905. *Journal for the History of Astronomy*, 1(1), 56–78. <https://doi.org/10.1177/002182867000100108>

Trahan, L. H., Stuebing, K. K., Fletcher, J. M., & Hiscock, M. (2014). The Flynn effect: a meta-analysis. *Psychological bulletin*, 140(5), 1332–1360. <https://doi.org/10.1037/a0037173>

Woodley, M. A. (2012). A life history model of the Lynn-Flynn effect. *Personality and Individual Differences*, 53, 152–156. <https://doi.org/10.1016/j.paid.2011.03.028>