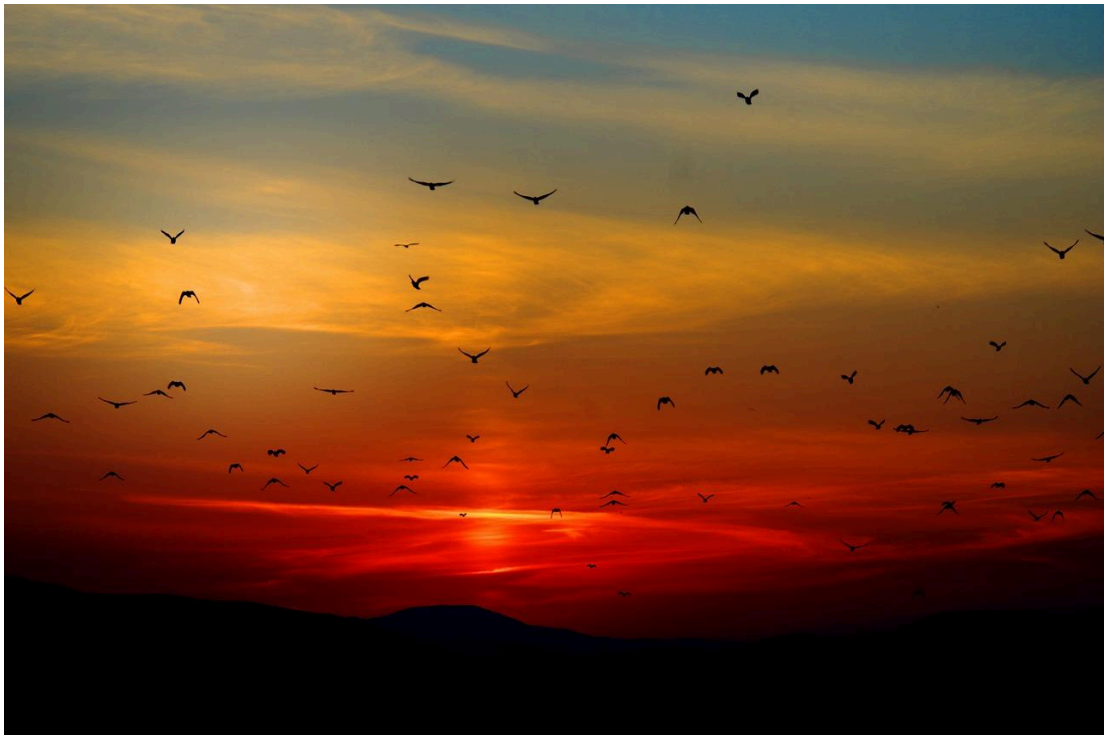


Noesis

The Journal of the Mega Society

Issue #213, May 2024

About the Mega Society

The Mega Society was founded by Dr. Ronald K. Hoeflin in 1982. The 606 Society (6 in 10^6), founded by Christopher Harding, was incorporated into the new society and those with IQ scores on the Langdon Adult Intelligence Test (LAIT) of 173 or higher were also invited to join. (The LAIT qualifying score was subsequently raised to 175; official scoring of the LAIT terminated at the end of 1993, after the test was compromised.) A number of different tests were accepted by 606 and during the first few years of the Mega Society's existence. Later, the LAIT and Dr. Hoeflin's Mega Test became the sole official entrance tests, by majority vote of the membership. Then, Dr. Hoeflin's Titan Test was added. (The Mega Test and Titan Test were also compromised, so Mega Test scores after 1994 and Titan Test scores after August 31st, 2020 are currently not accepted; the Mega and Titan cutoff is 43 - but either the LAIT cutoff or the cutoff on Dr. Ronald K. Hoeflin's tests will need to be changed, as they are not equivalent.) The Mega Society now accepts qualifying scores on The Hoeflin Power Test and on The Ultra Test. Both tests are still being scored. The Mega Society publishes this irregularly-timed journal.

Answer sheets for The Hoeflin Power Test and The Ultra Test can be emailed to ultrapowertest@gmail.com; the scoring fee for each test is $10 USD, payable via Stripe.

https://megasociety.org/#admission

The society also has a (low-traffic) members-only email list. Mega members, please contact one of the Mega Society officers to be added to the list.

For more background on Mega, please refer to Darryl Miyaguchi's "A Short (and Bloody) History of the High-IQ Societies" —

http://miyaguchi.4sigma.org/BloodyHistory/history.html

—and the official (designed) Mega Society page,

http://www.megasociety.org/

Noesis is the journal of the Mega Society, an organization whose members are selected by means of high-range intelligence tests.

Brian Wiksell (P.O. Box 366, Solana Beach, CA 92075) is the Administrator of the Mega Society. Inquiries regarding membership should be directed to Brian Wiksell at the aforementioned P.O. Box or the following email address: bwiksell@megasociety.org

Opinions expressed in these pages are those of individuals, not of *Noesis* or the Mega Society.

Noesis #213, May 2024

Editorial

Richard May, Ken Shea

Welcome back. The current issue explores aesthetics, natural law, metaphysics, giftedness, intelligence testing, political correctness, AI, age/cognition, and so-called conspiracy theories.

In the leadoff position, Chris Cole proposes a unifying principle for assessing “The Greatest of All Time (GoAT)” aesthetically, with examples spanning a few modes.

Next, Ian Williams Goddard applies “Frameworks for Respiratory Ethics” to the pandemic, employing natural ethics as a guide. Does *ahimsa* (non-violence) broadly apply?

Then, professor of psychology and author of *Malignant Self-Love: Narcissism Revisited*, joins Scott Douglas Jacobsen for a discussion of the interplay between physics, mysticism, the philosophy of science, and metaphysics in “Physics or Metaphysics?”

Sam Vaknin and Scott Douglas Jacobsen proceed apace with “Interview on Giftedness and IQ” by way of considering intelligence, IQ, giftedness, and genius.

In “Conversation with Bob Williams on General Intelligence Now” frequent contributor and Triple Nine Society member Bob Williams returns to survey various definitions of intelligence, *g*, the apparent Flynn Effect, detrimental effects of political correctness, giftedness, and related topics.

In “Conversation with Bob Williams on Practical and Impractical Intelligence Testing” Bob Williams reflects on retirement, the perceived difference between physics and differential psychology, observed correlations between the brain’s size/structure and intelligence, dysgenics, attempts at assessing IQ beyond the four-sigma level, Spearman’s Law of Diminishing Returns, Network Neuroscience Theory, and so-called multiple intelligences.

Next, researcher Stephen P. Badham presents results from multiple studies (“The Older Population is More Cognitively Able Than in the Past and Age-related Deficits in Cognition are Diminishing Over Time”) exploring how factors such as improved access to education, healthcare, and nutrition could be mitigating age-related cognitive decline globally.

Then, Ken Shea’s “Life’s Rich Pageant: John Updike’s Pulitzer Prize-Winning ‘Rabbit’ Tetralogy” contemplates a decades-running novelistic endeavor spanning the latter half of the last century.

Finally, May-Tzu, leaving no stone unturned, boldly explores “Unidentified Anomalous Geese (UAG) … and other conspiracy theories” in hopes of cracking the ultimate code.

Original submissions continue to be welcomed. Thank you to all of the current issue’s contributors. The next issue of *Noesis* will be published in December of 2024.

'And I understood that all these materials for a work of literature were simply my past life.'

-Marcel Proust, *Time Regained*

Contents

About the Mega Society		2
Editorial		3
Greatest of All Time (GoAT)	Chris Cole	5
Frameworks for Respiratory Ethics	Ian Williams Goddard	7
Physics or Metaphysics?	Sam Vaknin & Scott Douglas Jacobsen	15
Interview on Giftedness and IQ	Sam Vaknin & Scott Douglas Jacobsen	18
Conversation with Bob Williams on General Intelligence Now	Bob Williams & Scott Douglas Jacobsen	21
Conversation with Bob Williams on Practical and Impractical Intelligence Testing	Bob Williams & Scott Douglas Jacobsen	31
The Older Population is More Cognitively Able Than in the Past and Age-related Deficits in Cognition are Diminishing Over Time	Stephen P. Badham	48
Life's Rich Pageant: John Updike's Pulitzer Prize-Winning 'Rabbit' Tetralogy	Ken Shea	70
Unidentified Anomalous Geese (UAG) … and other conspiracy theories	May-Tzu	78

Greatest of All Time (GoAT)

Chris Cole

For any given form of art, the GoAT (Greatest of All Time) is the work that incorporates all possible instances of the art form. How is that possible? It must demonstrate how to create such art. Thus it latently contains all instances of the art form.

GoAT novel: A novel about making a novel

In Search of Lost Time - Marcel Proust

Reviewed by Jay Cantor:

“It’s a first person story of a character who shares the same name as the author and has a similar biography (though he is different in many ways, too). The book is an account of many things - the social world of the upper class in France, among them - but is primarily the story of how the character Marcel, who has the ambition to become the writer, learns how he might remember and describe his work, how he might tell the story you just read about how he became a writer.”

GoAT movie: a movie about making a movie

2001: A Space Odyssey - Stanley Kubrick

The star-child represents the work of art.

Reviewed by Christopher Nolan:

(https://variety.com/2018/artisans/production/christopher-nolan-2001-a-space-odyssey-1202811669/amp/)

“One of the reasons I hold Kubrick in such high esteem is I think he’s inimitable. Any time you see a filmmaker drawing too specifically [from Kubrick], it seems to not work. It seems to be self-conscious. He’s calm in the way he presents information, and there’s a simplicity and a discipline to his work that I think any filmmaker would aspire to. But he’s working on a plane far above the rest of us. That’s inspiring, but it’s also daunting.”

GoAT video game: a game about making a game

The Stanley Parable - Davey Wreden

Review by James Stephanie Sterling:

(https://www.destructoid.com/reviews/review-the-stanley-parable/)

“How do you review a game like The Stanley Parable? To describe any one part of it is to risk its ruination. To detail what it has to say about game design, the illusion of choice, and the psychology of the gamer is to tell you too much.

Comparisons, too, are going to be woefully inadequate. Perhaps its closest cousin would be Dear Esther, but where Dear Esther wastes the form of interactive entertainment, Stanley Parable uses and then subverts it. Where so many games that aspire to be more than games end up less than any form of art, Stanley Parable strives, and then succeeds, to be every game ever created. Even so, holding the game to the standards of any other title is simply not going to be correct.

So, how do you review what has become known as The Stanley Parable HD, the full-scale reimagining of one of the most intriguing mods available online? How do you discuss it, analyze it, and recommend it?

That’s quite simple.

You don’t.”

Frameworks for Respiratory Ethics

Ian Williams Goddard

Prior to Covid-19, who could have believed one day we would all be walking around wearing face masks? Masking for disease control in public had not been seen in most nations since the 1918 influenza pandemic. Then overnight there we were, all masking. My intuition was that being mandated to act so as to reduce the risk that I might spread a contagious novel virus that could kill millions before vaccines were available was compatible with natural ethics, which is to say, it was compatible with what one ought to do under the extraordinary circumstances even in the absence of a mandate. However, how do I know that intuition was reasonable?

Contemplating that question during the pandemic led me to the following ethical frameworks for respiratory hygiene.

Common Air as Common Property

My moral intuitions informed me that a body of terrestrial atmosphere, or air, serving as a common reservoir of life-sustaining oxygen is a *common property* of all its breathers. In other words, each breather has an equal co-ownership stake in that air. Why? Because each requires air to survive and it is naturally free worldwide. Therefore, if you have a right to life, you have a right to breathe that free air. Every breather, therefore, has an equal co-ownership stake in, and thus an equal right to access, the air around them.

Following that reasoning, if a body of air is collectively owned by all its breathers, no single breather has a unilateral right to pollute it without the consent of the others. The air you exhale is likely to be inhaled by others given that they must inhale or die. So, for example, it is unethical to exhale cigarette smoke into the air that others breathe if they do not consent to inhaling it.

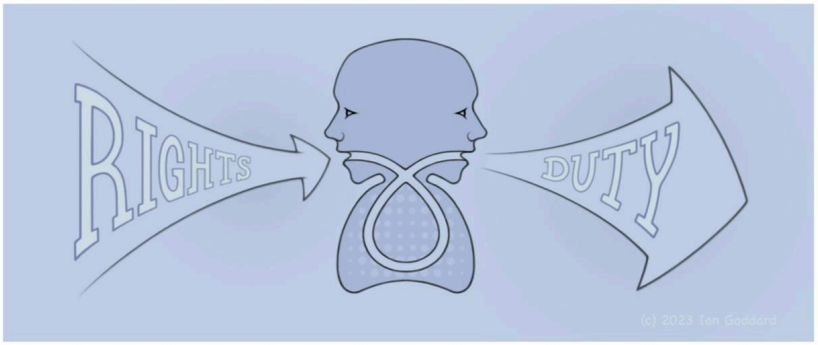


Figure 1 depicts the personal dynamics of respiratory ethics wherein everyone has a *right* to not *inhale* pollution (harmful substances) emitted by other people into their common air. Reciprocally, everyone has a *duty* to not *exhale* harmful substances into the common air they share with others.

Likewise, it is unethical to knowingly exhale harmful infectious germs into the air that others breathe if they do not consent to inhaling them. This is an ethical framework of shared air that outputs moral rights and duties of breathers, as depicted in Figure 1.

There is, however, a more fundamental substrate for respiratory ethics than the common property of air. Polluting the air that others breathe is an ethical issue primarily because it is *harmful*. Even if we concluded that shared air can be privately owned, that would not entitle an owner of a body of shared air to physically harm others by polluting it any more than a landlord is entitled to physically harm their tenants, which of course they are not so entitled. Therefore, common ownership of shared air is a sufficient but not necessary condition for respiratory ethics. A more necessary fundamental substrate is *harm* and the ethical duty to not cause physical harm to the bodies of other people.

Non-harm Undergirds Respiratory Ethics

From the Golden Rule to the Vedic principle of *ahimsa* to the Hippocratic Oath to the Lockean harm principle, the ethical precept that we should not cause harm to others, especially physical harm, endures from antiquity to this day. It therefore seems safe to say that knowingly causing physical harm to others with no justifiable reason is universally recognized as morally wrong, as something that you *ought* not do. It is even recognized that accidentally causing physical harm to others incurs a moral obligation to compensate those injured thereby. One of the ways you can cause physical harm to the bodies of other people is by being a vector of an infectious disease and thereby spreading your infection to others.

Thanks to the discovery that contagious illnesses are spread by germs, we were able to determine measures to mitigate those illnesses. Such measures included water treatment, refuse management, food inspection and personal measures for respiratory hygiene such as not spitting in public places, washing hands and covering coughs and sneezes. After the discovery of germs, respiratory hygiene evolved from measures for personal health to include measures to prevent spreading harmful germs to others. And not just to prevent the harm an infectee might directly spread to others, but also harms that those whom they infected could in turn spread to others and so forth unto untold subsequent transmissions and magnitudes of aggregate harm.

Given the non-harm principle, this argument follows for every ‘you’ out there.

> **Premise**: In accord with the non-harm principle, you should take precautions to prevent yourself from causing bodily physical harm to others.
>
> **Conclusion**: Given that being a vector of an infectious pathogen can cause others to suffer mild to catastrophic bodily physical harm, according to the Premise, you should take precautions to prevent yourself from becoming such a vector.

This conclusion comports with social norms wherein since the dawn of germ theory we have been admonished to cover coughs and sneezes and stay home when sick. This norm is especially meaningful because the aggregate harm an infectee can cause by failing to prevent spreading their infection has an unknowable upward limit. The still-accruing aggregate harm unleashed by the one person who was the index case of Covid-19 includes millions of deaths, immeasurable suffering and economic loss worldwide to date. Even if you infect only one other person with the flu, that could lead to enormous downstream harm including sick leaves, medical costs and deaths among untold numbers of infections stemming from yours. When you have a respiratory infection, you have the power to prevent potentially massive aggregate human suffering by following the argument specified above.

A Meta-framework for Risk-mitigation Frameworks

An important feature of risk mitigation is that risk varies over time. For example, the risk of harm someone poses to others increases when they are infectious with a contagious pathogen. It also increases when the likelihood that you are asymptomatically infectious rises during an epidemic wave. As the risk of harm we pose to others increases, our ethical obligation to prevent ourselves from causing that harm increases proportionally and in turn obliges us to use more effective measures to prevent that elevated risk. In this way, a duty to prevent harm and the measures taken to do so scale with changing levels of risk. This proportionality is depicted in Figure 2, wherein a three-level ordering simplifies what exists in experience as a smooth gradient of increasing and conversely decreasing degrees of obligation.

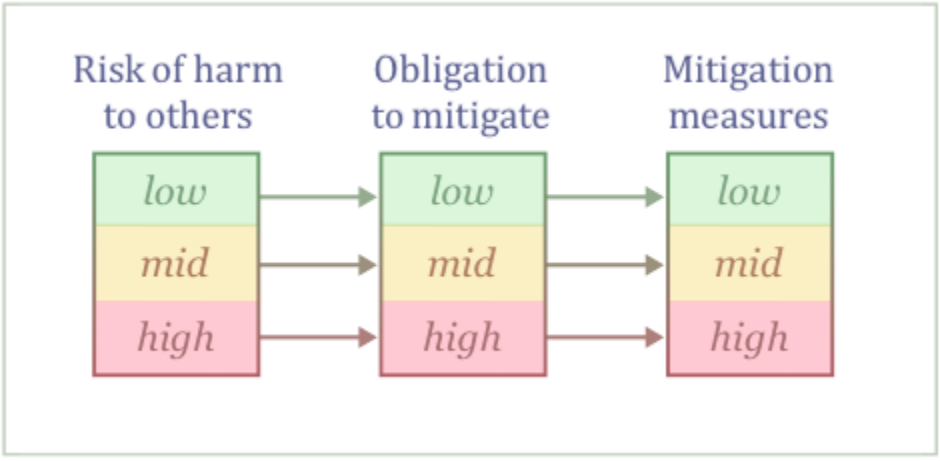


Figure 2 shows a proportional relation between magnitudes of risk (left), ethical obligation to mitigate that risk (center) and mitigative measures taken (right). When the risk you pose to others increases, your ethical obligation to mitigate that risk increases along with the mitigative measures you should take.

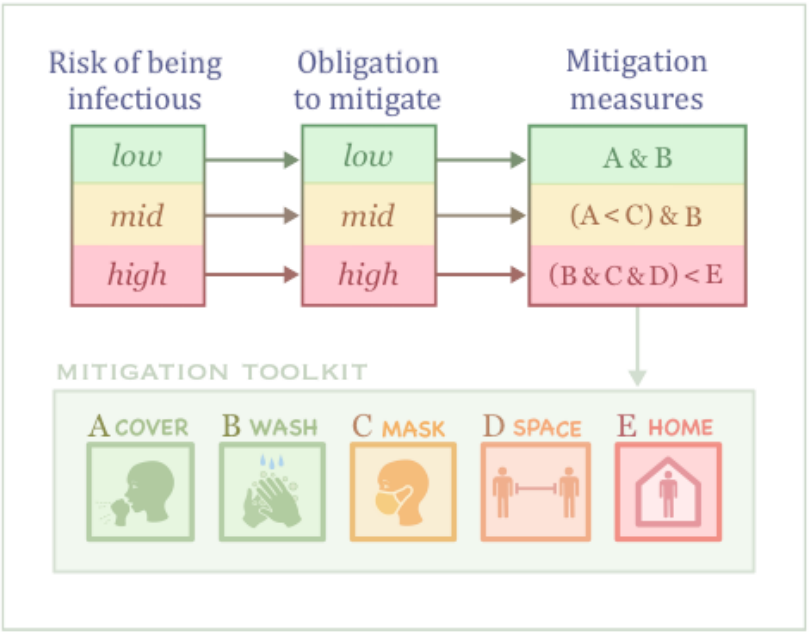


Figure 3 instantiates the proportionality framework in Figure 2 into the risk domain of respiratory hygiene. The right-hand column prescribes a mitigation protocol that scales from the easiest measures to measures that are more effective and layered, all derived from a mitigation toolkit. The < operator is a weighted inclusive or such that *x* < *y* means: *x* and/or preferably *y*.

This proportionality framework reveals the logical structure of respiratory ethics as they already exist, which is to say it is a *descriptive* more than *prescriptive* framework. Figure 3 instantiates the framework in Figure 2 into the risk domain of respiratory hygiene and adds a mitigation toolkit for this domain. Other mitigation measures in this domain, such as ventilation and vaccination, could be included as well. This toolkit is not intended to be comprehensive but rather an example of how mitigative measures are scaled and layered in proportion to changing levels of risk.

This proportionality framework is robust across domains of risk. For example, public officials have an ethical duty to protect the citizenry from harm posed by criminal convicts whose risk to others varies from lower or higher. That gradient of risk entails a proportional gradient of duty to mitigate that risk with a proportional gradient of measures from probation to incarceration in lower to higher security facilities. This proportionality is illustrated in Figure 4.

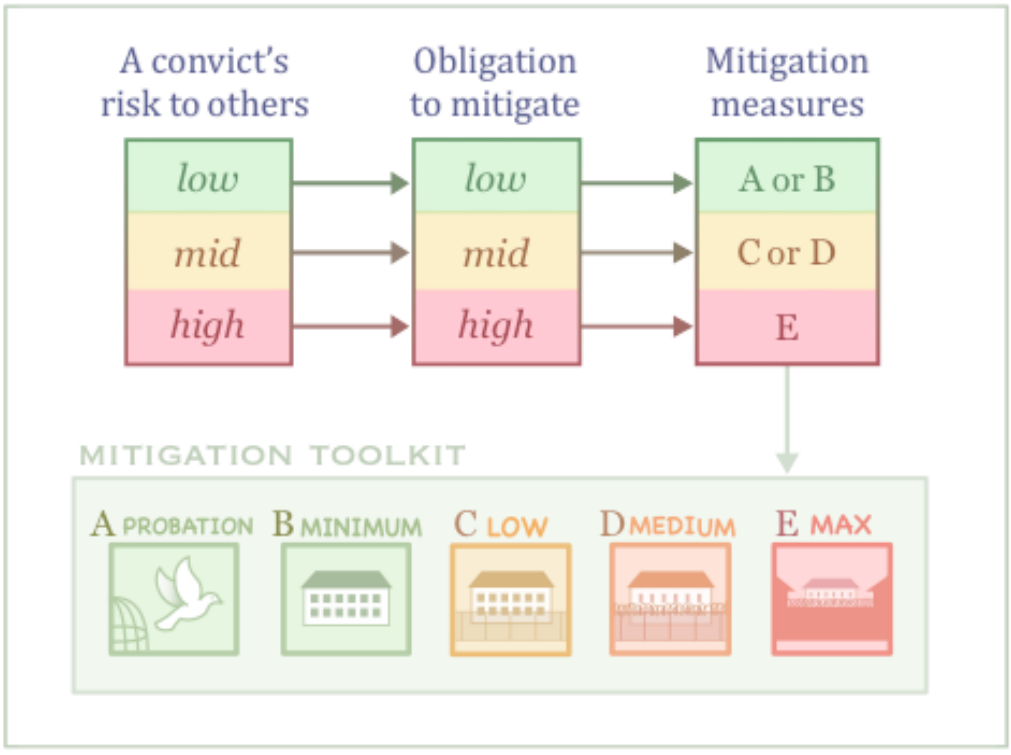


Figure 4 instantiates the proportionality framework into the risk domain of criminal justice. The risk of harm convicts pose to others varies from low to high placing correspondingly low to high obligation upon public officials to mitigate that risk with correspondingly less or more effective incarcerative measures. While the lowest-risk convicts may be free on probation, higher risk posed to others is mitigated in the United States with the four levels of prison security depicted here.

Let us consider a third instantiation of the proportionality framework, this time into the risk domain of firearm ownership. If you own a firearm, you incur an ethical duty to mitigate the risk of harm it poses to others, which includes a duty to store it safely. Your stored firearm poses a high risk of harm to others if children or someone with harmful intent finds it. A stored firearm owned by someone who lives alone probably poses less risk to others than one owned by someone who lives with others especially kids. This gradient of risk and duty entails a corresponding gradient of mitigative measures as illustrated in Figure 5, which instantiates the proportionality framework into the domain of firearms-risk management.

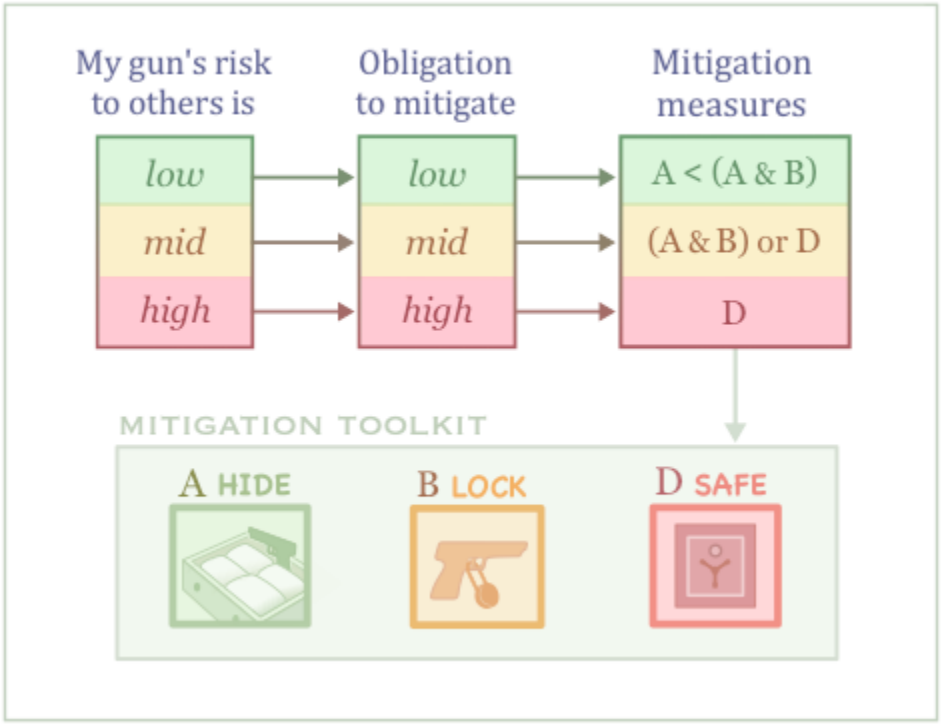


Figure 5 instantiates the framework into the risk domain of firearms safety. Storage of a firearm poses a risk to others depending on the likelihood that someone, such as a child or criminal, can get hold of it and thereby harm themself or others. The greater that risk, the greater your obligation to store your firearm safely and to do so with more effective measures.

These examples in Figures 4 and 5 show that the respiratory ethics framework in Figure 3 is an instance of the meta-framework of risk mitigation depicted in Figure 2. It is intuitively logical that if we should not cause physical harm to others, then our ethical obligation to take actions required to prevent causing physical harm to others should scale proportionally with the magnitude of the risk of such harm we pose to others. If we pose little risk of physical harm to others, our obligation to mitigate that risk is low and so at most we need only take simple and easy measures to mitigate that low risk. However, as our risk of causing physical harm to others increases, it is both intuitive and logical that our obligation to mitigate that risk increases proportionally, which in turn entails we ought to apply proportionally more vigorous measures to mitigate that greater risk.

Other Variables Affecting Respiratory Ethics

The framework for respiratory ethics in Figure 3 accommodates the basic variables regarding the likelihood that you are infectious. However, in various contexts other variables can influence decision making about respiratory ethics, including, (1) the vulnerability of others around you to severe infection outcomes, (2) the prevalence of masking among others around you, and (3) the presence of masking policies. These additional three variables are considered in Figure 6, forming three decision matrices. Let us consider each contextual case.

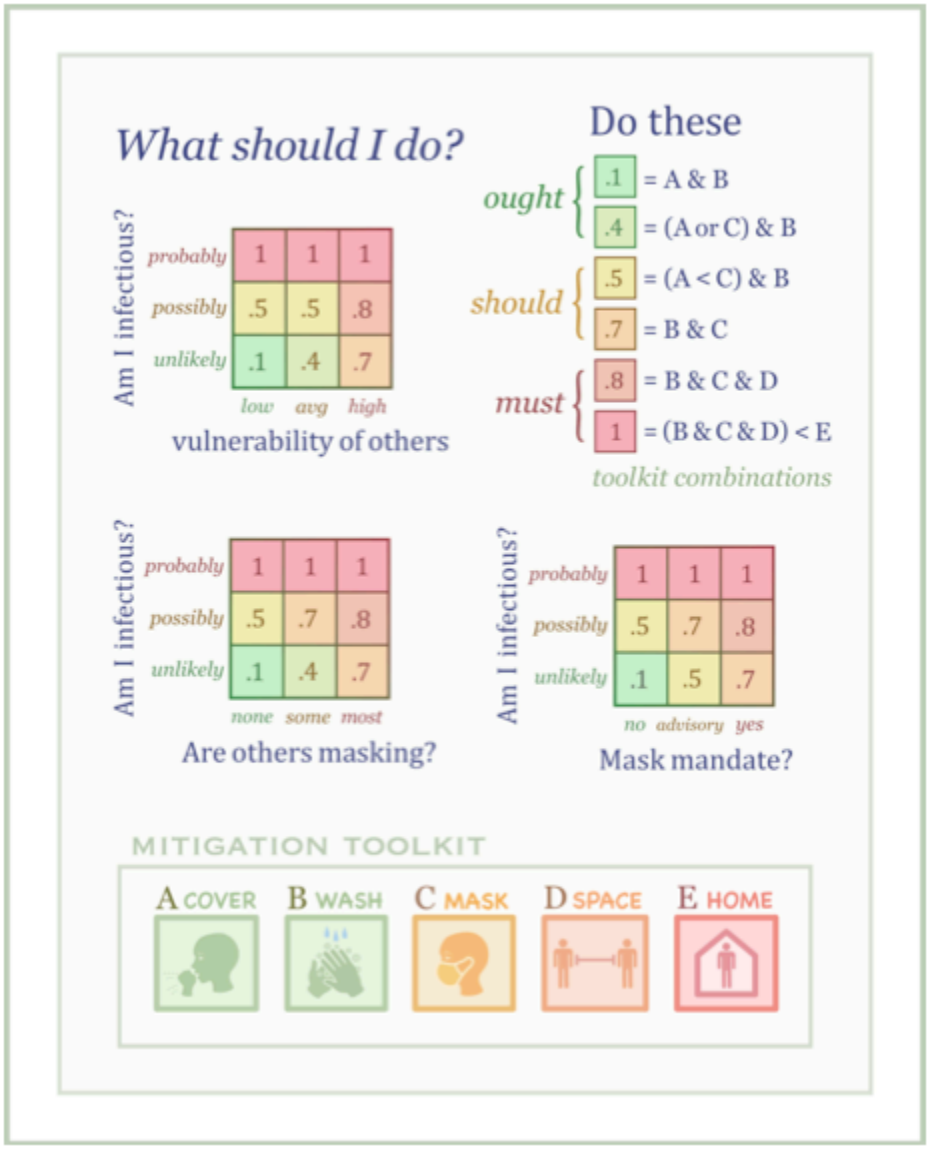


Figure 6 extends the decision model in Figure 3 to provide ethical answers for twenty seven two-variable contexts. The first model considers the vulnerability of others. The second considers the preponderance of others masking as a gauge of *common consent* for or against unmitigated spread of respiratory effluent into a shared breathing space. The third considers whether a masking policy exists and how that may affect your ethical decision making.

In the first case, suppose you are planning to visit a nursing home whose residents have a higher risk of suffering significant harm from infections. In that case the risk you pose to these specific other people around you has increased compared to the risk you pose to the average person you encounter outside the nursing home in the unlikely, albeit not impossible, chance that you are asymptomatic and infectious. Now you have a reason to decide that you should wear a mask to protect nursing-home residents even if you believe you are probably not infectious.

In the second case, if most people around you are masking, it may be reasonable to infer that most of them prefer to not breathe your respiratory effluent. You may, therefore, decide to recognize and respect that implicit common consent and don a mask even if you believe it is unlikely that you are infectious.

In the third case, if there is an institutional masking advisory or mandate — for example, if a nursing home requires visitors to mask — you may choose to wear a mask even if you feel fine. And you may choose to do so not merely because you wish to follow rules, but because the existence of such policies probably corresponds to actual higher levels of risk you pose to others under the particular circumstances that brought about out such policies.

Is Ethical Obligation Variable?

It might seem that the magnitude of ethical obligation is actually the same irrespective of the magnitude of risk. For example, you should mitigate low risk with easy measures and you should mitigate high risk with more efficacious measures. In each case you *should* mitigate risk and that can seem like the same magnitude of obligation applies to different levels of risk and the only stratification of levels at hand are the levels of measures you should take, but not levels of ‘shouldness’. If there are no levels of shouldness, the proportional framework in Figures 2-5 is inaccurate. However, it is easy to disabuse oneself of that impression by considering the different moral *blameworthiness* for failing to mitigate high risk versus low risk. If it is more unethical to fail to mitigate high risk, then ethical obligation itself does in fact range from less to more.

Consider for example person A and person B who both know they should store their handgun in a secure location. And suppose A lives alone and is the only person in her house and she leaves her handgun unattended out on a table in her house, while B is a father living with his young children who have friends over and he leaves his handgun unattended out on a table in his house. Are both due the same magnitude of moral censure for failure to mitigate risk? Of course not! Person B is obviously due far harsher moral reprimand for his failure to mitigate the risk his firearm poses to others. Yes, both *should* mitigate risk (what seems the same), but person B is uniquely under a higher magnitude of ‘shouldness’. Person B must mitigate the risk they pose to others whereas A merely *ought* to.

Let us consider another example. Suppose person A and person B both know they should cover coughs and sneezes. And suppose that A feels fine and by all indications is probably not infectious, while B is obviously sick and he believes he’s probably infectious. Now suppose both fail to cover a sneeze in close proximity to others. Are both worthy of equal moral censure? No! Obviously person B is most worthy of reprimand because there is reason to believe the risk he poses to others is higher. That differential magnitude of moral reprimand is precisely a differential magnitude of shouldness, confirming a gradient nature of ethical obligation.

So upon reflection it is easy to see that the ethical duty to mitigate risk does in fact vary from being less to more obligatory corresponding proportionally to lower and higher risk. So the variation in degrees of ethical obligation in the middle set of the proportionality framework in Figures 2-5 is an accurate model of how ethical obligation exists in real life.

Summary

Motivated by ethical questions arising from the Covid-19 pandemic, the respiratory ethics framework presented here addresses questions about why, how and when we should modify our practices of respiratory hygiene, as for example by donning a face mask. This framework is grounded in the moral principles of nonmaleficence, or non-harm, beneficence and respect for personal autonomy. From the nonmaleficence principle we derive a duty to engage in beneficent action for the benefit of others by acting to protect them from our own respiratory infections. Because this framework relies on a person's own assessment and mitigation of their risk to others rather than on top-down enforcement, it respects personal autonomy.

From a proportionality framework depicted in Figure 2, wherein increasing risk of harm to others entails increasing ethical obligations to mitigate that risk with increasingly efficacious measures, the framework's basic decision model is instantiated in Figure 3. The framework is then extended into two-dimensional decision matrices in Figure 6 to guide ethical decision making in more complex contexts. Figures 4 and 5 show how the generalized proportionality framework in Figure 2 can be instantiated in other domains of risk for decision making therein.

After contemplating the ethics of the airborne transmission of risk during the pandemic, I had a paper presenting my conclusions published in the *Journal of Health Ethics*. Given the professional-ethics focus of the medical-ethics journals, my paper is unique in presenting a personal-ethics framework applicable to everyone. From my survey of the literature, this might be the only formal ethical framework for airborne infectious respiratory disease.

Physics or Metaphysics?

Sam Vaknin & Scott Douglas Jacobsen

Scott Douglas Jacobsen: What would define a comprehensive physics?

Professor Sam Vaknin: Ostensibly, physics is the science of studying reality. But, in reality, physics is a form of mysticism, it is where alchemy used to be a few centuries ago.

In alchemy, there was a preoccupation with language and a belief in a universal, invariant truth which would endow the practitioner with godlike powers. All alchemists believed that it was only a question of time before we attain this truth.

Similarly, physics is a self-contained, self-referential language. There are debates on how to use this language and on how to interpret its elements. But there is a broad agreement on its grammar and syntax.

Jacobsen: What would define a complete metaphysics?

Vaknin: Metaphysics deals with concepts that underlie reality.

Jacobsen: What would relate these two universes of discourse in the aforementioned definitions?

Vaknin: Physicists still believe that physics is asymptotic to the truth, that we are making progress towards an objective, invariant, immutable, indisputable verity. This is, of course, mysticism, not science.

Scientific theories are not about reality but about other scientific theories and about themselves, a discourse that spans generations and contrasts with previous ways of thinking even as it generates falsifiable testable hypotheses.

Theories are allegories, metaphors, analogies, glorified literature using a highly structured language known as mathematics. Scientific theories are descriptive, predictive, and wrong: after all, all past theories have been falsified.

Physics is an extension of metaphysics. We must revert to philosophy and metaphysics, our roots.

All scientific theories are fundamentally metaphysical. Examples: evolution is founded on teleology (the accepted truth that organisms wish to survive) and SRT (special relativity theory)

emanates from the separateness of observers from the observed (which was proven wrong on the micro level).

The philosophy of science is a fancy rebranding of metaphysics. Its main tenet, falsifiability, is tautological (we can falsify only scientific theories which are the only theories that are falsifiable).

There is an age-old confusion between language and truth. For example: the solutions to an equation (language elements) are considered to be true and real. The very reliance on language is metaphysical because it assumes that language correlates with reality or can be perfectly mapped onto it.

But how many of our assumptions about language (such as axioms in mathematics) are real or true? We can study language only with a meta-language and this results in an infinite regression. So, there is no way to prove or to ascertain the validity or the relevance of a language. It takes a leap of faith.

Science is, therefore, a faith-based system that is helpful to survival (akin to religion). The core percepts are metaphysical, non-provable, they require a leap of faith. Physicists arbitrarily assume the validity and power of mathematics and the existence of reality - both are metaphysical assumptions which are unprovable, axiomatic, and not derivable.

I am a believer in physics, but I am not a naïve believer: it works, so I believe in it, but I am aware of my own irrationality. Reason is not primary, faith is.

Jacobsen: What was metaphysics in the past?

Vaknin: What we today call science. The study of both the essence of the world and of what makes it tick.

[Editors' Note: https://www.newworldencyclopedia.org/entry/Philosophy_of_nature]

Jacobsen: What has been the origin and evolution of physics into the present?

Vaknin: There were two major revolutions in the history of physics and its divorce from metaphysics: Descartes' and Bohr's.

At some point in time, we started to believe that observer and observed are two separate, unrelated systems. In the twentieth century, we gave up on any pretension and attempt to capture the quiddity of the world or even to merely describe it. Instead, we settled on an instrumental version of physics: if it works, it is futile to inquire why and how it works. Quantum mechanics is a prime example of this blindfolded approach.

Jacobsen: How is physics beginning to turn into, or make a circumlocution back to, metaphysics, and into an evolution of “uber-metaphysics” - even mysticism? What are the dangers - let’s say - to clarity of concepts and thought in turns, some of them, to mysticism now?

Vaknin: The minute you let language dictate your view of reality (because it is “self-efficacious”), you abandon the latter. The formalism rules and the procedures for manipulating its symbols become the laws of physics or of nature.

The problem with this kind of detour is that, as Gödel has observed, formal-logical systems are incomplete or inconsistent and give rise to “hallucinations” (witness the recent debacle with artificial intelligence).

[Editors’ Note: https://plato.stanford.edu/entries/goedel-incompleteness/]

So, we end up further away from a true understanding of reality as we descend into arcane solipsistic sophistry.

Jacobsen: How is physics, in some sense, like a highly formalized structure of literature?

Vaknin: I don’t think that it is like literature. Physics is not merely descriptive. It doesn’t confine itself to a taxonomy of the codification of experience. It aspires to decipher reality even as it translates into technology: tools to alter our environments, near and far.

Jacobsen: All scientific theories in the past have been proven wrong via experiment or come to inconsistent findings with Nature and the predictions of the theory. What will happen to the current set of theories, most likely?

Vaknin: The same fate awaits the current crop of dominant theories.

Jacobsen: If physics, currently, is rootless or physicists - as a category - have ‘forgotten’ their foundations in metaphysics as physics being a derivation from metaphysics, what is a necessary bridge to bring the roots back to soil for physicists - and for the structured narrative knowledge pool called physics to become crisp again?

Vaknin: Philosophy and logic. These should become mandatory studies. These disciplines are indispensable in the evolution of critical thinking and the generation of testable hypotheses.

Jacobsen: Thank you for the opportunity and your time, Professor Vaknin.

Vaknin: Pleasure, as always.

Interview on Giftedness and IQ

Sam Vaknin and Scott Douglas Jacobsen

Scott Douglas Jacobsen: You have been measured three times with a high IQ, an understatement. With the IQ scores of 185 at age 9, 180 in the army at age 25, and 190 in prison at age 35, presumably on a standard deviation of 15.

What was the reaction of family, friends, peers, community, even the psychometricians or psychologists administering the tests each time?

Professor Shmuel "Sam" Vaknin: First, let me clarify that any result above 160 (some say, 140) is not normatively validated: it is rather arbitrary and meaningless because there are so few people to compare with (the sample is way too small). Matrix IQ tests are better at validating higher results, though.

Everyone always loathed me. I am a sadist, so from a very early age, I have leveraged my IQ to taunt people, hold them in contempt, and humiliate them. This did not endear obnoxious me to anyone. My own teachers sought to undermine my academic career, peers shunned or attempted to bully me (they failed), my mother detested me, my father pendulated between being awe-struck and being repelled by me. Both my parents beat me to an inch of my life every single day for 12 years.

Jacobsen: To you, as a scientific person, what defines intelligence?

Vaknin: Anything that endows an individual with a comparative advantage at performing a complex task constitutes intelligence. In this sense, viruses reify intelligence, they are intelligent. Human intelligence, though, is versatile and the tasks are usually far more complex than anything a virus might need to tackle.

Jacobsen: What defines IQ or Intelligence Quotient?

Vaknin: The ability to perform a set of mostly – but not only - analytical assignments corresponding to an age-appropriate average. So, if a 10 year old copes well with the tasks that are the bread and butter of an 18 years old, he scores 180 IQ.

IQ measures an exceedingly narrow set of skills and mental functions. There are many types of intelligence – for example: musical intelligence – not captured by any IQ test.

Jacobsen: What defines giftedness, to you? Even though formal definitions exist.

Vaknin: Giftedness resembles autism very much: it is the ability to accomplish tasks inordinately well or fast by focusing on them to the exclusion of all else and by mobilizing all the mental resources at the disposal of the gifted person.

Obviously, people gravitate to what they do well. Gifted people have certain propensities and talents to start with and these probably reflect brain abnormalities of one kind or another.

Jacobsen: Interrelating the previous three questions, what separates intelligence from IQ from giftedness, i.e., separates each from one another?

Vaknin: IQ is a narrow measure of highly specific types of intelligence and is not necessarily related to giftedness. Gifted people invest themselves with a laser-focus to effect change in their environment conducive to the speedy completion of highly specific tasks.

Jacobsen: What defines genius?

Vaknin: Genius is the ability to discern two things: 1. What is missing (lacunas) 2. Synoptic connections.

The genius surveys the world and completes it by conjuring up novelty (i.e., by creating). S/he also spots hidden relatedness between ostensibly disparate phenomena or data.

Jacobsen: How does genius differentiate from intelligence, IQ, and giftedness?

Vaknin: A genius can have an average IQ or even not be analytically very intelligent (not be an intellectual). Some craftsmen are geniuses. Musicians, athletes, even politicians.

Jacobsen: What happens to most prodigies, or adults with exceptionally, profoundly, or unmeasurably high IQ?

Vaknin: A majority of them end badly. IQ is a good predictor of academic accomplishments, but not much else. Character, upbringing, mental illness, genetics, nurture, the environment (including the physical environment), sexual and romantic history matter much more than IQ.

Many “geniuses” with a high IQ (Mensa types) are dysfunctional and deficient when it comes to life, intimacy, relationships, and social skills. Additionally, as Eysenck had correctly observed, creativity is often linked to psychoticism.

Jacobsen: What are the optimal things for raising gifted children and prodigies, and for resuscitating drifting adults with exceptionally, profoundly, or unmeasurably high IQ, if at all possible, to productive and healthy lives?

Vaknin: All interventions are somewhat effective only during childhood and adolescence, up to age 21. Afterwards, it is an uphill battle.

The most crucial thing is to never remove the gifted child from his peer group (as was done to me). I am also dead set against academic shortcuts.

The gifted child should follow the same path as everybody else but feed his voracious mind with extracurricular enrichment programs and materials.

Jacobsen: Who seem like the greatest geniuses in history to you?

Vaknin: The usual suspects: Einstein, Newton, Freud, da Vinci, other polymaths who had upended every discipline or field that they had turned their scintillating minds to.

Jacobsen: Thank you for the opportunity and your time, Professor Vaknin.

Vaknin: The opportunity is all mine.

Conversation with Bob Williams on General Intelligence Now

Bob Williams & Scott Douglas Jacobsen

Abstract

Bob Williams is a Member of the Triple Nine Society, Mensa International, and the International Society for Philosophical Enquiry. He discusses: a background in nuclear physics, interest in intelligence, and the transformation of Fort Langley due to the influence of Trinity Western University; retirement in 1996 as a pivotal moment for deeper exploration of human intelligence, access to scientific resources and the internet for furthering studies, and involvement with the International Society for Intelligence Research since 2003; shifts to definition of intelligence, critique of the APA's definition and suggestion of alternatives, emphasis on the importance of psychometric *g* and the role of genetics and environment in intelligence; addresses misinterpretation of the Flynn Effect, explanation of its non-relation to genuine intelligence increases and citation of examples of IQ decline in developed nations, challenge to the notion of environmental improvements enhancing intelligence; touches on political and social ramifications of intelligence research, impact of "woke" culture on academic freedom and dismantling of programs for gifted students, sharing of personal anecdotes from interactions with notable researchers; comments on enduring relevance of *The Bell Curve*, contributions to the field, and global variability of the Flynn Effect, concluding with insights into genetics of intelligence and challenges facing contemporary intelligence research.

Keywords: Cultural Shifts, Dysgenics, Education, Environmental Factors, Flynn Effect, Genetics, Heredity, Intelligence, IQ Tests, Nutrition, Psychometric *g*, Research, Retirement, Social Intelligence, Technology.

Scott Douglas Jacobsen: Today, we're back with Mr. Bob Williams, retired super smart guy! Former nuclear physicist and participant in interviews on IQ and intelligence in *In-Sight Publishing* and republished in *Noesis*: The Journal of the Mega Society. Most of my best friends as a 13-year-old into the present have been near-retired or retired people, I grew in an artsy, intellectual town called "the village" also known as Fort Langley. It is different now. The Evangelical Christians from Trinity Western University have, more or less, made the place wealthier, tiny bit snooty, and much more glossy. Yet, they call the place, still, "the village." To each their own, Fort Langley, when I grew up, was a retirement place, a quietude. So, retired people are the best people in my opinion! Do you find yourself having more time to pursue interests in retirement?

Bob Williams: I retired when I was young, in 1996, and regard that move to be one of the best of my life. Since I have a lot of interests, having more time has enabled me to spend more of it with these interests and to both enjoy them and to improve my expertise in them. My interest in human intelligence began in the early '90s, when I was working in Washington, DC (Department of Energy – Senior Technical Advisor). Having a scientific library there (this was when we still used MicroFiche for research) gave me access to some papers that I would have otherwise found difficult to obtain. When I retired, I had more time to study this new passion, which was aided by increasing electronic access to resources and ultimately to the newly available internet. I joined the International Society for Intelligence Research (ISIR) in 2003 and started attending its conferences in 2004. This opened a new world of access… directly to the people who were writing the papers and books I had been reading.

Jacobsen: The American Psychological Association in "**Intelligence**" defines intelligence, in an adaptation from the Encyclopedia of Psychology, as follows:

"Intelligence refers to intellectual functioning.

Intelligence quotients, or IQ tests, compare your performance with other people your age who take the same test. These tests don't measure all kinds of intelligence, however. For example, such tests can't identify differences in social intelligence, the expertise people bring to their interactions with others.

There are also generational differences in the population as a whole. Better nutrition, more education, and other factors have resulted in IQ improvements for each generation."

Given their use of the Encyclopedia of Psychology, I will use this as a resource, too. Jensen is deceased; Flynn is dead. Many larger names in intelligence research's history are passed. I do not know if significant changes or developments have occurred within the field of research of general intelligence. However, the institutions devoted to psychology have been changing norms and mores, which, in turn, adapts the empirical frameworks' orientation: what is emphasized more, what is emphasized less. Does this definition seem adequate for a beginning definition of intelligence?

Williams: Before I get to your question near the end, I think it is worth arguing a bit with the APA definition of intelligence. It is not totally off, but I don't think it is as good as these:

The best definition:

Intelligence = psychometric *g*

The most cited:

Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings–"catching on," "making sense" of things, or "figuring out" what to do.

Source: Linda Gottfredson – Mainstream Science on Intelligence; *The Wall Street Journal*; December 13, 1994 — signed by 52 intelligence scholars.

[Editors' Note: https://www1.udel.edu/educ/gottfredson/reprints/1997mainstream.pdf]

My favorite is Carl Bereiter's clever definition:

"Intelligence is what you use when you don't know what to do."

The problem with the APA definition is that it tries to downplay the importance of intelligence and then adds the misleading two sentences at the end. This has been a trend of woke people before the word identified socialism and extreme anti-science rhetoric. Nutrition has not been a factor in developed nations for a long time. The brain needs iron, iodine, and folate to develop properly. These are present in the diets of all developed nations and all but the most backward others. Education does not change real intelligence, it simply provides us with the tools we need to do various cognitive tasks. Intelligence is determined by the DNA we inherit and may be reduced by encounters with the environment (disease, toxins, and head trauma).

Throughout any discussions of intelligence, we must understand that intelligence is about biology and that it is fairly equated to psychometric *g*. Researchers refer to this as a Jensen Effect, meaning that if something is not observed as a change in *g*, it is not a Jensen Effect and is not about the essence of intelligence. We will get to a lot of this in relation to the Flynn Effect.

The assumption relating to IQ improvements for each generation is at odds with a substantial amount of data showing that real intelligence has been declining for a long time in virtually all developed nations. The dysgenic effect on intelligence has been extensively reported in scholarly papers and books. Here are three examples of books reporting it:

Herrnstein, R. J., & Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.

At Our Wits' End: Why We're Becoming Less Intelligent and What It Means for the Future, by E. A. Dutton & M. A. Woodley of Menie. Exeter, UK: Imprint Academic.

Lynn, R. (2011). *Dysgenics: Genetic Deterioration in Modern Populations* (revised ed.). London: Ulster Institute for Social Research.

The APA definition also wants us to buy into the Multiple Intelligences nonsense that was successfully pushed on laymen and has stuck like molasses. We only need to consider *g* (or, to a lesser extent, the residuals of broad abilities, after *g* is factored out) when we are discussing intelligence. Psychometric *g* accounts for essentially all of the predictive validity of IQ tests and it is only because those tests can be used as proxies for *g* that they have any real utility.

[https://www.niu.edu/citl/resources/guides/instructional-guide/gardners-theory-of-multiple-intelligences.shtml -Editors' Note]

It is misleading to imply intelligence enhancing environmental factors that simply do not exist. Researchers have not yet found a single thing in the environment that increases intelligence.

For at least the past 5 years, we have had some open discussions (ISIR conferences) of the importance of finding a way to increase intelligence. Despite our world-class neurologists, geneticists, and psychologists, none claim any means of increasing *g*, but all agree that it is a desirable goal. Now that we finally know what defines intelligence, the prospects of doing it via genetics seems unlikely until amazing new technologies appear.

The actual question, which I have somewhat evaded, is about changing norms, mores, and the APA definition. My view on the definition is hopefully clear. Norms and mores have become more antagonistic towards researchers, who have had the courage to deal with the relatively short list of deadly topics: differences in intelligence between breeding groups and the sexes, and to a lesser extent the heritability of intelligence. I know researchers who are totally afraid of being connected with any aspect of these three topics. They have seen careers ruined, people losing their jobs, physical threats, physical attacks, vandalism, denied promotions, and speakers being invited to universities only to be shouted down, followed by police escorts to protect them from mobs. Yes, it is serious and nasty.

One of the consequences of the woke culture is that schools for bright students have been abolished or crippled to such an extent that they have been reduced to ordinary schools with names that suggest otherwise. Thomas Jefferson High School for Science and Technology has been repeatedly named by *U.S. News and World Report* as the number 1 high school in the United States. It used testing as a major part of its selection process. The school board eventually reached a woke majority and proceeded to disallow testing for admission. The stated reason was that the board noticed that 68 – 70% of the students were Asian and most of the rest were Whites. So now, students are admitted on the basis of skin color, instead of intelligence. New York effectively has done the same thing, not to one extraordinary school, but to all gifted programs. For more information than you would ever want to read, see this search result:

https://www.bing.com/search?q=new+york+eliminates+gifted+education

This same process is apparently being repeated in other woke states. Bright students have become an embarrassment to school boards. At TJHSST (see above), National Merit finalists were not notified of their success until it was past time for them to apply for related scholarships and to their accomplishment on college applications. The school administration said that they did not want those who were not selected to have their feelings hurt. Then it was found that 14 high schools in Fairfax County did exactly the same thing and that this had been ongoing for ten years! The real reason behind the withholding of the notifications was that most (or all) of the finalists were Asian or White. That is where our norms and mores have gone.

Jacobsen: Implicitly, this definition refers to the Flynn Effect, not coined by James Flynn, but **Richard Herrnstein** and **Charles Murray** in their 1994 book The Bell Curve: Intelligence and Class Structure in American Life. How did this mistaken identity of the title, the Flynn Effect, get the attribution?

Williams: I will paste in the introduction to my paper on this subject:

The secular rise in IQ scores appeared unexpectedly and has defied explanation. Smith (1942) recorded a gain (in Honolulu) over a 14-year span. Later, Tuddenham (1948) found an increased intelligence when he compared inductee scores for the U.S. Army from World War I and World War II and proposed that the gains might be due to increased familiarity with tests; public health and nutrition; and education [the gains from 1932 to 1943 were 4.4 points per decade.]. He cited a high correlation (about .75) between years of education and the Army Alpha and Wells Alpha tests that he was studying.

The secular gain remained relatively dormant until it was rediscovered by Lynn (1982) while working on a comparison of Japanese and U.S. data. It was then rediscovered again, using American data, by Flynn (1984a,b). The raw score gains did not have a name until Herrnstein & Murray (1994) coined the term Flynn effect in their book *The Bell Curve* (p. 307). Some researchers choose to refer to the secular gain as the Lynn–Flynn effect, or use an uppercase FL (FLynn effect) for the obvious reason that they feel Lynn has been somewhat slighted by not including his name.

Source: Williams, R. L. (2013). Overview of the Flynn effect. *Intelligence*, 41, 753-764.

Jacobsen: Flynn, in my interviews with him, firmly believed Murray was not a racist. He was the liberal counterpart in this general intelligence and IQ debate. He described the entrance into the debate and the academic as one motivated by liberal leanings. Murray is conservative. Whether consciously or not, with this as a political affiliation, this would affect research questions for Murray, eventually, and the orientation within the research chosen. In this case, the research on IQ. Thus, the split between the liberal orientations and conservative frames on then IQ debates generically tends towards environmentalist versus hereditarian. Although, as Noam Chomsky has noted, it's trivial to say heredity plays a role in traits. It's like claiming something was the result of evolution in biological systems, including spandrels, because everything in biology is a result of evolution writ large: All forms of selection. Therefore, if someone claims a trait isn't hereditary to a minimum degree – a non-zero level, then they're not part of the serious discussion on attempts to pin down a) a definition of human intelligence and b) measurements for this definition in order to create a functional and repeatedly measurable psychological construct. As the counterpart to Murray, it seems natural to assume an ad hominem, especially given the current intellectual climate. Yet, he does not do this. He knows Murray very well as another researcher looking to conclude the opposite of Murray. Furthermore, and to reiterate the point, near the end of his life, he did not see Murray as a racist. What do you make of this claim against Murray?

Williams: I have had the good fortune of knowing both (Flynn and Murray) and to chat with them, sometimes for long times, at the conferences we attended. I have distinct impressions of both and will share my thoughts. I first met Flynn in 2007 in Madrid. I found him to be warm and pleasant to talk to, while behaving differently when he was in front of our group. He had a booming voice and used it to silence people by literally drowning them out. He had a lot of exchanges with Jensen over many years, with both parties remaining respectful of the other. In these exchanges, it is my belief that Jensen was consistently right and Flynn was not. Flynn was totally honest about how his political beliefs came into play, both in relation to his

employment woes and in his beliefs about intelligence. Jensen, as a true opposite, looked at data and nothing else. He reported what he found in data and allowed no other factors to distort what was measured and (usually) replicated.

Flynn was respected by lots of big name researchers. I felt that this was not justified and once wrote something to that effect in response to a comment on Roberto Colom's blog. I was surprised when Roberto asked me if I would write an explanation of my comment for publication on his blog; I did. Those who read Spanish can find my reply here:

https://robertocolom.wordpress.com/2015/11/18/cambios-seculares-en-las-puntuaciones-de-los-test-de-ci-explicaciones-y-omisiones-de-james-r-flynn-por-robert-l-williams

For those who would like to see the original reply (in English), use this link:

https://www.dropbox.com/s/6negb8rno2lvl9b/Flynn%27s%20explanations%20and%20omissions%20%28Bob%20Williams%29.pdf?dl=0

In my reply, I discussed some of my thoughts on how Flynn approached various topics. He avoided the use of unambiguous terminology, avoided topics that would not support his positions, and even tried to support his ideas by inventing scenarios (magic multipliers, as reported with Dickens) that are not derived from data and which are at odds with the findings of researchers over the past 50 years.

Below are some comments from Linda Gottfredson that are parallel to my impressions.

'Flynn's Fallacies

With characteristic understatement, Flynn says that everything became clear to him when he awoke from "the spell of *g*" (pp. 41-42). The reader, feeling afloat in a rolling sea of images and warm words, might ask whether he succeeds only by loosing himself from the bonds of evidence and logic. More troubling, his core argument rests on logical fallacies that profoundly misinterpret the evidence. I describe three below. To be fair, they are among the common fallacies bedeviling debates over intelligence testing, and most reflect a failure to appreciate the inherent limitations of psychological tests, including tests of intelligence.'

Source: Shattering Logic to Explain the Flynn Effect; Linda S. Gottfredson • November 8, 2007 • Cato Unbound.

Murray is more like Jensen, in that he makes his arguments based on data, not politics. Like Flynn, I found Charles to be friendly and very bright. In any technical argument that one might imagine between them, I would expect the sound, accurate, and realistic argument to come from Murray.

Things have changed drastically over the past decade. We used to get updates from Robert Plomin about every 2 years (at ISIR conferences), concerning the search of genes relating to IQ. I recall that he once told us that the SNP chips that they were using could not possibly fail to detect a gene with as much as a 1% effect size – yet there was nothing. Fortunately, genome wide association studies arrived and the missing links appeared. Researchers found that

intelligence is defined by tens of thousands of single nucleotide polymorphisms, not by individual genes. When I asked James Lee (one of the pioneers in this work) how many SNPs were geneticists estimating as defining intelligence, he told me the range was from 10,000 to 40,000. When the genomic data set reached over 1.1 million genomes, researchers found 1,271 SNPs that were associated with high intelligence. The average effect size of these SNPs is 0.01%. Together they can account for 10% of the variance in intelligence

Effects as tiny as these can only be seen when GWA studies reach sample sizes of tens of thousands of cases for disorders such as schizophrenia, or hundreds of thousands of unselected individuals for dimensions like educational outcomes. As GWA studies reached these daunting demands for statistical power, they struck gold. But what GWA studies found was gold dust, not nuggets. Each speck of gold was not worth much, but scooping up handfuls of gold dust made it possible to predict genetic propensities of individuals.

Robert Plomin – *Blueprint: How DNA Makes Us Who We Are*, Penguin Books Ltd., 2018, ISBN 9780241282076 [Editors' Note: http://tankona.free.fr/plomin2018.pdf]

Since individual DNA is set at the moment of conception, estimates of IQ can be made before birth [Using DNA to predict intelligence; Sophie von Stumm, Robert Plomin; *Intelligence* 86 (2021) 101530], during life, or thousands of years after death. [See Intelligence Trends in Ancient Rome: The Rise and Fall of Roman Polygenic Scores; Davide Piffer, Edward Dutton, Emil O. W. Kirkegaard; *OpenPsych* July 2023; DOI: 10.26775/OP.2023.07.21]

Anyone who argues the environmentalist side of the old argument is not living in the present. That story has been told to such an extent that we can safely say that there is not even a scent left to sniff. No environmental effects have been shown to increase *g*. Even the home environment has been shown to have essentially no impact on intelligence (based on MZA twin studies and adoption studies, including interracial adoption studies). [MZA = monozygotic twins reared apart]. But this goes much further. Stephen Pinker's very long book *The Blank Slate*, is an overkill showing that even other behavioral traits are primarily associated with the non-shared environment, not the shared (family) environment.

The last time I saw Jim Flynn was in 2017. Here is one of the pictures I took when he was addressing ISIR:

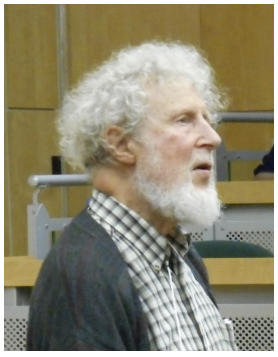


Image Credit: Bob Williams

Jacobsen: The basic premise in the argument against *The Bell Curve* has been one-sided: Charles Murray is a racist. Let's say, that's so. Assume the premise, does this have any impact on the foundational presentation of the work?

Williams: *The Bell Curve* was understated and bulletproof. Herrnstein and Murray went to great lengths to not overstate anything and to document everything they discussed in terms of how intelligence relates to life outcomes. They also wrote personal interpretations of how intelligence would impact our lives in the future and offered ideas as to how to deal with such outcomes. It was always clear when they were giving opinions.

Today we have the benefit of major breakthroughs in brain imaging and genetics. Many issues that were not fully settled in 1994 are no longer subject to argument. Today we have a massive increase in worldwide intelligence studies that are so detailed that it is possible to map IQ variations within nations. In 1994 there were few studies of remote and underdeveloped nations, but that is no longer true. *The Bell Curve* remains as probably the best and broadest study of how intelligence shows up in the lives of different populations. The idea of first showing 12 chapters of data for non-Latino whites, then showing that the same effects are seen in blacks was brilliant.

Jacobsen: Herrnstein was the math guy. Murray is the social stuff guy. With Herrnstein dead so early as the text gained traction, did this impact the proper interpretation of the full statistical analysis of the work?

Williams: It is unlikely that Herrnstein's death had any impact on the book. Writing began in spring of 1990. Herrnstein died on September 13, 1994 (less than 2 weeks before publication). Herrnstein was diagnosed with lung cancer in June 1994. I don't know when he stopped working on the book, but it is fair to say that virtually all of the composition work was done well before he died.

In 2019 ISIR awarded Murray with the Lifetime Achievement Award. During his related speech, he mentioned that, while at MIT, he took every course on data analysis that was offered by the university. He had already decided what he wanted to do as a career and it was not political science. I have no idea how the work was split between Herrnstein and Murray, but I expect that a significant amount of the analytical work was done by Murray.

As many readers here know, Murray has addressed a number of topics in his books and columns. One that is related to *The Bell Curve* is *Facing Reality* (2021). I was impressed with his invention of an analytical method to measure eminence–used in *Human Accomplishment* (2003). He demonstrated that it was accurate by benchmarking the methodology against two sports that have massive amounts of quantitative measures of performance (baseball and golf).

Jacobsen: Is the Flynn Effect continuing or declining, or stagnating globally? My understanding: In some sectors of the world, it is continuing, while, in others, it is stagnating or declining. All at variable rates.

Williams: Yes, you are right. I think it may be helpful to list a number of salient points that apply to the Flynn Effect.

- The FE is not a Jensen Effect. It is not on *g* and, therefore, is not related to real intelligence. It is possible to select a cause that should be *g* loaded, but those have not been shown to actually apply. So, we must allow for the possibility that small Jensen Effects will be found in some places and times.
- At the present time, some nations are experiencing gains in IQ test scores; some are finding that their scores are in decline; and others are seeing no changes.
- At any time, when a FE is observed, it does not impact broad and narrow abilities equally. Some may be increasing while others are declining. When the FE was mostly associated with score increases, the gains were more prominent in abstract reasoning test items, while academic test items were decreasing.
- In some nations, there have been score increases, followed by stability, followed by score decreases. There is no evidence that the people in these nations showed increases in real intelligence during positive FE changes nor did they become duller as negative FE changes were found.
- Negative FEs have been reported in Norway, Denmark, Britain, Netherlands, Finland, France, and Estonia. The IQ decline rates, per decade, range from 1.35 to 8.4 IQ points. [See E. Dutton, et al./*Intelligence* 59 (2016) 163-169]
- The FE has been reported in preschool children, thereby eliminating at least those data from school-related causes.
- Some studies have found that the FE was stronger in the low IQ part of the IQ spectrum. Other studies found it mostly in the high-IQ range. And other studies found that it was equally evident in all ranges. I think that these inconsistencies are important because they point to artifacts and not group-level changes.
- Jensen commented that the definitive test of whether FE gains are hollow or not is to apply the predictive bias test. This means that two points in time would be compared on the basis of an external criterion (real world measurement, such as school grades). If the FE gains are hollow, the later time point would show underprediction, relative to the earlier time. This assumes that the later group has not been re-normed. In actual practice tests are periodically re-normed so that the mean remains at 100. The result of this re-centering is that the tests maintain their predictive validity, indicating that the FE gains are indeed hollow. If the gains were real and the tests were re-normed, people at a given IQ would be getting smarter and this would show up in the predictive validity. [Jensen, A. R. (1998). *The g factor: The Science of Mental Ability*. Westport, CT: Praeger.]

- Brand, C. (1996). *The g Factor: General Intelligence and Its Implications*. Chichester, England: Wiley [The book was withdrawn by Wiley after it was released. The reason was that it accurately addressed differences in the IQs of blacks and whites.] In this book, he noted that a probable cause of the FE was increased guessing. This is now known as the Brand Effect and has been documented in detail from Estonian data that covered 72 years. The Brand Effect can make score gains appear to load on *g*, when they do not. This happens because the most *g* loaded test items are the most difficult for low *g* persons, so they have more guessing and more gains.
- Another indication that FE gains are artifacts was shown by A. Beaujean, who scored National Longitudinal Survey of Youth data using both classical test theory and item response theory. When the superior IRT was used, the gains vanished in some cases and halved in others. This is entirely due to an external artifact and has nothing to do with intelligence.
- Rushton used principal components analysis to show the independence of the FE from known genetic effects. The data showed that the IQ gains on the WISC-R and WISC-III form a cluster. This means that the secular trend is a reliable phenomenon. This cluster is independent of the cluster formed by racial differences (shown by many replications to be differences in *g*), inbreeding depression scores (purely genetic), and *g* factor loadings. The secular increase is, therefore, unrelated to *g* and other heritable measures.

Conversation with Bob Williams on Practical and Impractical Intelligence Testing

Bob Williams & Scott Douglas Jacobsen

Abstract

Bob Williams is a Member of the Triple Nine Society, Mensa International, and the International Society for Philosophical Enquiry. He discusses: satisfactory retirement in 1996; how standardized tests were not widely utilized for nuclear physics job admissions; microfiche as a valuable research tool; entering workforce in 1966 without testing; transition from male-dominated colleges to coeducation; early '90s intelligence research material; Richard Lynn's work in Mensa Research Journal; influential books on intelligence research; statistical methods for high sigma tests facing challenges; challenges to psychometric *g* including alternative intelligence models; Network Neuroscience Theory exploring brain networks' role in intelligence; intelligence decline trends observed in developed nations; statistical methods not applicable in intelligence studies; the validity of high sigma IQ tests; constructing culture-fair tests for high sigma ranges facing practical and theoretical challenges; AI advancements and intelligence measurement; DNA analysis and intelligence estimation; AI conversational agents estimating human intelligence; fear of controversy may hinder certain research topics; respect for disciplines may be affected by controversial research topics; unaided smart kids in education; "woke" in context of left-leaning educational policies; potential avenues for measurement, exploring animal studies and leveraging AI technologies; concept of "magic multipliers"; decoupling of familial environment (FE) from general intelligence (*g*); ethical considerations of reproductive technologies, particularly in context of assisted reproduction and genetic screening; potential development of artificial general intelligence (AGI) based on our understanding of brain structures and processes related to intelligence; and integration of modern network models with existing theories of intelligence, signaling potential direction for future research in this field.

Keywords: admissions, challenges, conferences, diffusion tensor imaging, intelligence, interviews, libraries, microfiche, myths, networks, psychometric *g*, research, standardized tests, statistics, twin studies.

Scott Douglas Jacobsen: How was the retirement in 1996? Were standardized tests of note utilized in admissions for particular jobs in workspaces requiring nuclear physics? I have used MicroFiche in some research at one of the libraries in a postsecondary institution here. It is still a good resource. I'm pro-MicroFiche, but a minority-user!

Bob Williams: I entered the workforce in 1966. There was no testing, just a face-to-face interview. The thing that is interesting (to me) about the outcome of this is that hiring people largely on the basis of the degrees they held resulted in a fairly homogeneous group of people who ranged from bright to very bright. In 1966 we were still in an era in which a much smaller fraction of men went to college/university and a still smaller fraction of women went. Of the women who did attend college, most were in colleges for women (including some very well-known schools with respected academics) or went to colleges for teachers, which was a subset of the former. By the time I retired women were a majority in some colleges and the colleges that previously admitted only men were open to women. I think by then colleges for women were admitting men and the real, women only, colleges were headed for change or closure.

I am surprised that MicroFiche still exists! I love being able to locate papers and books with a computer and often obtain the found document instantly by downloading it.

Jacobsen: The period between the 1990s and 2003/04 of joining and attending conferences of the International Society for Intelligence Research. What were the first realizations in this independent research for you?

Williams: Back then, good material was not only more difficult to find, but there was much less of it. In the early '90s I subscribed to the *Mensa Research Journal*. It was mostly filled with reprints from various sources, but occasionally had a direct submission. I recall seeing Richard Lynn's work there and reading about his ideas about the evolution of intelligence. They presented him with an award for his intelligence research contributions. At about that time, I joined the International Society for Philosophical Enquiry and met Miles Storfer. I bought his recently written book from him (he carried them around): Storfer, Miles D. (1990). *Intelligence and Giftedness: The Contributions of Heredity and Early Environment*. San Francisco, CA, US: Jossey-Bass. Then a big one arrived: Herrnstein, R. J., & Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press. By this time, I had found and read enough material that I already knew the material they reviewed, so the interesting part was the new analysis of the National Longitudinal Survey of Youth data. A few years later, the most cited book in the history of intelligence research publications arrived: Jensen, A. R. (1998). *The g Factor: The Science of Mental Ability*. Westport, CT: Praeger. I had already read some of Jensen's papers and some references to his work in various other sources. By the time I met Jensen in 2004, he had become my passive mentor.

My realizations were, first that I had to learn some statistical methods that I had not previously encountered, and second that the science of intelligence is inherently messy. Coming from a physics background, I was used to things being precisely measurable and repeatable. The niche of intelligence within differential psychology was much like mud wrestling. I quickly learned to appreciate the challenge of extracting meaning from data that was full of confounds. It is a fascinating challenge and I think it is rewarding, particularly when most of the real meat of the science is hidden to a much greater extent than happens in physics and chemistry.

In the innately fuzzy world of life sciences there are studies that we cannot do for social or practical reasons, but someone finds a brilliant way to extract the information from natural experiments. For example we cannot inflict a famine on an experimental group, but since real famines have happened (such as the Dutch famine during WW2), it is sometimes possible to find data that relates directly to those events. Besides the Dutch data, there was the interesting question of how to determine if head sizes had changed over time. If you want to consider a long time, direct measurements are impossible, unless they were performed and recorded (they were not). In this case, Rushton found Army data on the number of military helmets that were issued by size. Yes, he found an increase.

Jacobsen: Were there points of collaboration?

Williams: Yes, a few. Most of the material I published was solo, but there were a few papers where I was a coauthor. These were all publications in academic journals. I have published much more in the private journals *Noesis*, *Gift of Fire*, *Vidya*, and *Telicom*.

Jacobsen: Let's call this the exploratory years or something friendly like this, what were the major realizations upon entering the field at the time? What were the first myths dispelled?

Williams: I don't recall having heard and believed any of the many popular myths that persist about intelligence. There were lots of new things to learn that I had not previously encountered. Learning how the twin studies and adoption studies were conceived, executed, and reported was important and impressive. Both Robert Plomin and Thomas Bouchard initiated these somewhat challenging studies. I met Bouchard in 2004 and recall having asked him enough questions to have been a pest. He was very helpful in explaining things that few people understand. For example, I learned that it was true that twins have a statistically lower intelligence than singletons and that the issue of the heavier twin being more intelligent was true, but had been solved by prenatal care. I also learned that the attacks against some researchers were much worse than I imagined. Among those who really suffered (in the time frame you mentioned) were Nyborg and Brand, both of whom lost their jobs. Jensen took more flack than anyone, but he seemed unfazed by it. In fact, he told me to watch for the upcoming paper he did with Rushton. He said that he expected it would cause "quite a stir." [Rushton, J.P. and Jensen, A.R. (2005). Thirty Years of Research on Race Differences in Cognitive Ability. *Psychology, Public Policy, and Law*, Vol. 11, No. 2, 235-294.] After the paper came out, I asked him if there was any notable reaction to it. He said "no," and seemed disappointed. It led me to suspect that he was looking forward to another rant from the left, which did not happen.

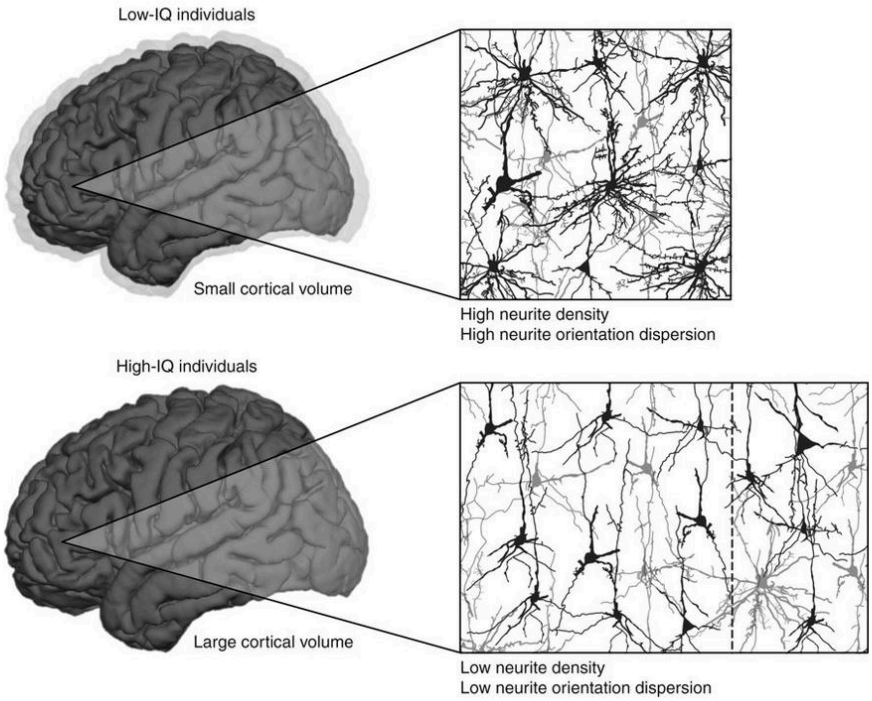
Jacobsen: Now, to those first realizations and myths taken away by truths, what ones have remained true?

Williams: I wish I had a list of such myths that involved me, but as I explained, there were none. I was disconnected from the field of intelligence research until my interest developed in the early '90s. When I became interested, I was lucky (or careful) to ease into the new field by following the real experts. The job was one of reading books and papers and those generally do not get far off target.

There was one common belief that was disproved to the surprise of everyone. One of the things that was consistently reported was the correlation between brain size and intelligence. When structural MRI became available, the correlation was found to be about r = 0.40. That was challenged by a meta-analysis that showed a somewhat smaller correlation coefficient, but then it was shown that the meta-analysis consisted of a large number of studies that used low-quality IQ tests. When only high quality tests were used, the old number turned out to be correct. But that was not the surprise. The surprise appeared in this paper:

Erhan Genç, et al. (2018) Diffusion markers of dendritic density and arborization in gray matter predict differences in intelligence; *Nature Communications* 9:1905. It can best be appreciated from this figure from the paper:

From: Diffusion markers of dendritic density and arborization in gray matter predict differences in intelligence



Schematic depiction of differences between low-IQ and high-IQ individuals with regard to brain volume, neurite density, and arborization of dendritic trees within the cortex. High-IQ individuals are likely to possess more cortical volume than low-IQ individuals, which is indicated by differently sized brains (left side) and differently sized panels showing exemplary magnifications of neuron and neurite microstructure (right side). The difference in cortical volume is highlighted by the shadow around the upper brain. Due to their larger cortices, it is conceivable that high-IQ individuals benefit from the processing power of additional neurons, which are marked by the dotted line in the lower panel. The cerebral cortex of high-IQ individuals is characterized by a low degree of neurite density and orientation dispersion, which is indicated by smaller and less ramified dendritic trees in the respective panel. Intellectual performance is likely to benefit from this kind of microstructural architecture since restricting synaptic connections to an efficient minimum facilitates the differentiation of signals from noise while saving network and energy resources. Neurons and neurites are depicted in black and gray to create a sense of depth. Please note, this depiction does not correspond to the actual magnitude of effect sizes reported in the study. For the purpose of an easier visual understanding, differences in both macrostructural and microstructural brain properties are highly accentuated

The explanation with the figure explains what was found. Genç was using diffusion tensor imaging for this work. I have had the great pleasure of getting to know him a bit. His most recent work combines brain imaging with polygenic stores.

Jacobsen: After the exploratory years and the interaction with individuals who wrote papers and books on the subject of intelligence, what first struck you about the professional community of intelligence researchers? Some see intelligence as the most important human trait.

Williams: Of course, intelligence is not only the most important human trait, but it is even more. Detterman expressed this perfectly:

Detterman, D. K. (2016). Was Intelligence necessary? *Intelligence*, 55.

“From very early, I was convinced that intelligence was the most important thing of all to understand, more important than the origin of the universe, more important than climate change, more important than curing cancer, more important than anything else. That is because human intelligence is our major adaptive function and only by optimizing it will we be able to save ourselves and other living things from ultimate destruction. It is as simple as that.”

As for the professional community, my impression was that the researchers were brighter than I expected and some were strong mathematicians (statistics). I also found that they were open to having a non-psychologist asking a lot of questions.

Jacobsen: What have been the most significant challenges to psychometric *g* as the definition of intelligence and as a psychological construct in the past? How have those been met with sufficient time and evidence?

Williams: The two well known challenges to *g* theory are Gardner’s multiple intelligence model and the emotional intelligence construct. Both are wildly popular among laymen and shunned by researchers. Both models contend that *g* theory is incorrect, but both are based on arguments in which *g* is present. For example, of the multiple intelligences claimed by Gardner, most are just statements of factors that are linked to the one and only *g*. Most book authors feel obligated to mention these models, then explain that they are not sound.

Jacobsen: What remain challenges to psychometric *g*?

Williams: There are some new models that are being discussed, but the literature that I have seen does not show a fully constructed model for any of them. Instead, they mention aspects of recent research that point to other model configurations. One of these is Network Neuroscience Theory. Relatively recent technologies, such as Diffusion Tensor Imaging, have made it possible to see and study brain networks. The characteristics of networks have shown that they are indicators of intelligence. The brain is, per this research, organized as a small-world network. This means that there are dense local networks (anatomically localized modules) that communicate with global networks. The modules have the advantage of close proximity within the small network, making them fast and efficient for related tasks.

[Editors' Note: https://www.cell.com/trends/cognitive-sciences/fulltext/S1364-6613(17)30221-8]

If the brain suffers focal injury, a module can alter its function to help compensate for lost modules in the damaged volume. This results in a more robust brain that can deal with trauma (to some extent).

Much of this is similar to the way we use networks for information movement between computers. It is my understanding that one of the difficulties is the wide range of structural differences between people. This is yet another demonstration of the messiness encountered

when trying to use neurological data statistically. It can be done, but requires a lot of separate observations, followed by good statistical analysis.

Anyone wanting to find and read material on this topic should begin by searching for papers by Aron K. Barbey. I have read his work for years and always found it to be outstanding.

[Editors' Note: https://psychology.unl.edu/aron-barbey

https://www.ncbi.nlm.nih.gov/myncbi/aron.barbey.1/bibliography/public/]

Jacobsen: Regarding "IQ improvements for each generation is at odds with a substantial amount of data showing that real intelligence has been declining for a long time in virtually all developed nations," what regions of the world have the strongest data and have the weakest data? What is the reason for the gap in depth of data?

Williams: Intelligence studies tend to start in Western Europe and North America, then are extended to other locations. One obvious reason for this is that there are more intelligence researchers in those two locations and it is much easier for them to do local studies. In the case of intelligence decline, there are multiple specifics that apply:

- The dysgenic effect was identified and described in *The Bell Curve* in 1994. Richard Lynn published a book on it in 2011, then Woodley and Dutton published another book (*Wits' End*) in 2018. *The Bell Curve* included only a small box on the topic, but the two books from Britain were focused on the decline. So, virtually all of the book-level work was British; this shows as a dominant factor in *Wits' End* (2018).

- Since the cause of the dysgenic effect is the negative correlation between IQ and fertility rate, the effect would be muted – probably to zero – in very low IQ nations and breeding groups (e.g. sub-Saharan Africa and Australian Aborigines).

- Since the effect size is small, it was easily masked by gains in the Flynn Effect (these are non-*g* artifacts). In order to study the actual changes over time, it is necessary to have data that goes back for over a century. Such data can be found in Britain and possibly a couple of other nations.

So, we cannot learn much about other nations, from direct data. These are discussed in *Wits' End*.

- The findings from the 1870s onward can be extrapolated to more recent reports, which now include essentially all developed nations.

Jacobsen: When there are gaps in data, are there statistical methods used to fill those gaps if they exist?

Williams: Not in this case. Per my comment above, the cause and effect has been established by data, largely from Britain, that goes back to Galton. Once the process has been shown by a variety of independent measures, we are left to accept the default hypothesis (that the same thing happens consistently) until something is identified to point to another outcome.

Jacobsen: If so, how do those statistical methods work?

Williams: I haven't seen any attempt to do more than demonstrate that the fertility rate is negatively correlated with IQ. There was some discussion of the role of increasing mutation load as a cause of the dysgenic effect. That topic died, probably due to the realization that tens of thousands of SNPs are the genetic basis of intelligence. With tiny effect sizes, accumulated mutations would take a very long time to show an effect.

One interesting and related area of research is the study of past civilizations by using polygenic scores. I have comments on this a few answers down. It may eventually be possible to use polygenic scores to make statistically reliable estimates of the changes in mean intelligence (for a given location) over time.

Jacobsen: What might be a hypothetical test with the ability to tap into 1-sigma and 6-sigma *g*? In theory, if the data continues to follow one after the other in a convergent direction, then we should have high-range tests with potentials for large properly controlled samples of the general population without compromises to the test. Chris Cole, a longstanding member of the Mega Society, and his team have been working for years on an adaptive test – cheat-resistant. David Redvaldsen's recent norming of the Mega Test and the Titan Test show test scores legitimate up the one-in-a-million level, but barely, and nowhere near many of the claimed scores of one-in-a-billion or more. Those remain false, but seemed true in an earlier time and the newer norms seem more reasonable given the newer spate of testing devoted, mostly independently, to the high-range. It is a testament of the contribution of Hoeflin to high-range testing to get above 4-sigma tests, but shy of 5-sigma.

Williams: There are two parts to my belief that measurements above 4 sigma are not informative: 1) norming is impractical; 2) the construct of intelligence and its measure (IQ) are difficult-to-impossible to defend. There is also a problem of demonstrating that high sigma tests can be compared over the same range.

As we all know, IQ is measured relative to a group of real people who are selected to statistically represent the full population. Typical professional IQ tests are designed to cover a range of ± 2.5 sigma, which is adequate to reach the 99th percentile. Some professional IQ tests (the WISC 4 & 5, Stanford-Binet 5, and DAS2 are the ones I am aware of) claim extended scales. They claim to use developmental markers instead of norming group data. Obviously,

this restricts the scales to children. The largest adult norming group I am aware of is 8,000 for the Woodcock-Johnson. Some tests have considerably smaller groups and presumably take a hit in the error bands for that reason. To test at 4 sigma, you would need over 31,000 people in the norming group in order to hopefully have one datum. It is easy to see that even at 4 sigma, the cost of dealing with a huge norming group would be prohibitive. The process effectively reaches an unbearable cost with very little return. [If Item Response Theory is used, norming is not required, but the need for a large reference group does not vanish.]

[Editors' Note: https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/wisc-v/wisc-v-technical-report-6-extended-norms.pdf]

Now, let's deal with construct validity and predictive validity. As we go beyond 4 sigma (and possibly before reaching it) we have to ask if the construct of IQ is the same as it is at lower levels. Because of Spearman's Law of Diminishing Returns (SLODR – if we accept it as fact), we expect that very high intelligence becomes heavily influenced by group factor residuals. [group factors = broad abilities, these are Stratum II in a three stratum model] In other words, the thing that we are doing at the usual levels is using a tool that had enough *g* variance that it can be used as a proxy for *g*, but SLODR tells us that *g* contributes less and less to the variance in intelligence as we move to high levels. Although the analogy is not perfect, you can think of this as being similar to the change of state of a solid as it is heated and becomes liquid, and then goes to a third state as a gas. The properties of the same element in each state cannot be meaningfully compared. In the case of measuring above 4 sigma, there is the likelihood that most of the variance is not *g* variance, so it is necessarily variance in the residuals of broad abilities, after *g* is factored out. Here we have a case of measuring where there is not a single *g* that is accounting for the inter-individual differences, so different people may score very high on any of the group factors.

In the CHC model, these factors should be present:

- Gc __ breadth and depth of acquired knowledge
- Gf __ fluid reasoning – reasoning, form of concepts, solve problems
- Gq __ quantitative knowledge
- Grw __ reading and writing ability
- Gsm __ short term memory
- Glr __ long term memory
- Gv __ visual processing – think and recall with visual patterns
- Ga __ auditory processing – process and discriminate speech sound
- Gs __ processing speed – clerical task speed

If *g* has already reached near saturation, factors such as Gf and Gc (top *g* loadings) probably will not turn out to be the source of most variance. Just guessing, I would expect Gq, Gv, and Ga might turn out to be dominant. If someone scores at a level taken to be at 5 sigma due to a very high Gq, would it make sense to say that he is equally smart as someone at the same 5 sigma level who made it on the basis of a high Ga? To me, the reason intelligence is meaningfully measurable over the usual range, is that it can ultimately be reduced to one single factor (*g*).

If we ignore all of the small details and have a test that specifies rarity up to 6 sigma, there must be real-world measures that confirm that the test is differentiating something that happens differently as a function of IQ in the very high range. The sorts of things that work in measurable ranges are similar to these: income, SES, job status, number of patents issued (engineers), age at tenure (professors), scientific publications, major awards*, having a role in work that is domain changing, etc. If outcomes cannot be statistically predicted for different levels (i.e.: 5 sigma vs 5.5 sigma) then the test is not meeting the requirement of predictive validity and must be classified as an ethereal exercise.

* Examples from the awards received by Feynman: Putnam Fellow · Nobel Prize in Physics · Albert Einstein Award · Oersted Medal · National Medal of Science for Physical Science · Foreign Member of the Royal Society.

Since I have already made this answer long, I will not expand much on the various other items that relate to difficulties in measuring above 4 sigma, but I will list some of the things that have to be resolved if a test is to be useful at any level:

- Is it invariant with respect to breeding groups, sex, and age?

- Is it properly and confidently age corrected so as to meet the definition of IQ? [I think this is an important one.]

- Is it subject to Flynn Effect artifacts? Are they properly handled?

- Is the *g* loading of the test known? [Requires testing a large group.]

- Is the reliability coefficient derived from sound measurement? Is it 0.90 or higher?

- Is construct validity established by comparison between its factorial structure and that of a major comprehensive test (WAIS or Woodcock-Johnson)?

- Are the broad ability factors balanced, so that the test is not unduly weighted by a small number of factors? [This impacts the factor loadings of the test.]

- Is the test administered by a qualified person (psychologist)? If not, how is the use of new and powerful artificial intelligence prevented?

[These and similar items were discussed in my article, High Range IQ Tests — Are They Psychometrically Sound? *Noesis* #207, February 2021.]

All of these things are difficult to satisfy and are usually quite costly. It may be impossible to actually demonstrate some, or most of these for ceilings above 4 sigma.

Jacobsen: How could we use techniques for translating regular gold-standard tests like the WAIS and SB to make culture fair tests up to a 6-sigma range?

Williams: Given my long answer (above), I believe that the problems I listed are unlikely to be resolved unless something startling appears from AI. The surprises that are coming from AI are more than a step-up, they are dramatic. The particular study that I think illustrates how AI can do things that were not only unexpected, but also not understood by researchers: Banerjee, I., Bhimireddy, A.R., Burns, J.L., Celi, L.A., Chen, L.C., Correa, R., Dullerud, N., Ghassemi, M., Huang, S.C., Kuo, P.C. and Lungren, M.P., 2021. Reading race: AI Recognises a Patient's Racial Identity in Medical Images. *arXiv* preprint *arXiv*:2107.10356.

This x-ray analysis, based on AI, demonstrates that something totally unforeseen might happen that changes how intelligence is best measured and understood. One area that I am watching is the analysis of genome-wide association studies, using AI.

Jacobsen: If *g* is largely innate while still susceptible to environmental blunting, can we estimate the contexts of *g* for ancient civilizations and peoples, as a general comparative metric in current times, so making a within-species general comparative metric across times? People likely encountered more bodily traumas and malnutrition in the past, for instance. Modern Western types, in most cases, tend to be well-fed, pampered, and comfortable in contrast with ancient humanity.

Williams: There is IQ work ongoing now, based on DNA samples from ancient groups. The first paper I encountered on this topic: Intelligence Trends in Ancient Rome: The Rise and Fall of Roman Polygenic Scores; Davide Piffer, Edward Dutton, Emil O. W. Kirkegaard;*OpenPsych* July 2023; DOI: 10.26775/OP.2023.07.21. There is a long video interview of Piffer by Kirkegaard that discusses this topic in depth. I assume readers can find it with a search engine. Piffer mentioned that DNA data is pouring in from various ancient groups and that there is ongoing work to analyze it via polygenic scores. There are some obvious limitations, such as not being able to identify insults to the DNA that might have reduced individual intelligence. As the sample sizes increase, the confidence levels of this work will improve, but even now, the results are useful in tracking intelligence over wide time intervals.

Jacobsen: In the future, could we use artificial intelligences mimicking various general levels of intelligence of people to do wordplay and that converse with human interlocutors to estimate *g*

in the tested human? It would be a step away from a direct brain scan estimate, but it would be cheaper and more output oriented.

Williams: I assume that AI will advance from the already impressive performance (certain applications) to reach levels that will be startling. AI should be able to learn from various data sets, such as the norming data for the Woodcock-Johnson that has been made available to researchers. It would seem to be a natural fit for the use of Item Response Theory. AI should be able to determine Item Characteristic Curves, or something similar, but which is developed from within the AI system. I wouldn't be surprised if it is eventually able to make good estimates of intelligence by simply examining discussions by various people, either in video or text format. We already do that when we watch someone who is either obviously dull or obviously brilliant. It would be interesting to see what a trained AI system would, perhaps in a few years from now, observe from videos of Joe Biden, Kamala Harris, Christopher Hitchens, and Sabine Hossenfelder.

Jacobsen: In theory, could we use such a system to establish what a human general intelligence – whatever the culture and native tongue – would likely produce as output in conversation if intent on showing the real general intelligence, even if we have not found such an individual through regular testing channels with a psychometrician? There is popular theatrical commentary on an LLM with an IQ of 155 for verbal intelligence. Stuff like this. However, I mean a real correlation matrix extended or extrapolating based on live human input and incredible amounts of data and deep learning, ANNs. So, "Human with a cognitive rarity of 1 in a 1,000 sounds like this on either side of the curve. 1 in 30,000 sounds like this. Therefore, based on these sophisticated algorithms and extrapolations, the 1-in-10,000,000 person should sound like this." It would reverse the sample size problem to an artificial sample size solution in a way. An artificial constellation of language used to determine where someone sits in cognitive rarity with the ANN constantly learning, improving with each additional human interlocutor. It would be a narrow band artificial intelligence with this specific purpose, especially good with the large amount of correlation with *g* and verbal ability, e.g., like Hogwarts's Sorting Hat minus the magic.

Williams: Yes, I agree with the likelihood that AI will be able to match behavior or language to a given specification. It would be the reverse direction of the prior question. I think that it would have a lot of leeway for a given level of intelligence, since we already know that you can name a percentile and find a wide range of behaviors at that level. AI should be able to match the intended IQs of fictional characters that are described as input.

I have doubts that this sort of thing would retain meaning when the end of the range of the definition of intelligence (pre my prior comments) is reached.

Jacobsen: Do newer generations of intelligence researchers feel a tinge of fear for asking particular research questions when seasoned researchers encounter "careers ruined, people losing their jobs, physical threats, physical attacks, vandalism, denied promotions"? I sense a chill among both conservatives and liberals, oddly less amongst centrists, in sociopolitical

contexts. Both use cancellation as a tactic. That's not new. Lots of us have experienced it. I don't care about it much, personally. The advancement of knowledge is the key part. For the advancement of a field with key impacts, it raises legitimate, serious concerns about the advancement of research in the terms of the potential for rapid developments for benefit for humanity as a whole, especially the floor of societies who benefit from smart, dedicated people with ethics bent towards general humanitarian efforts. Identification and nurturance efforts matter. You noted this in the last part.

Williams: I see two things happening. The first is that some researchers are fearful of discussing anything that might lead to a hot topic or even allow someone to claim that they have commented on one. The fear is what I assume went on when the Roman Catholic church punished Galileo in 1633. Other scientists could see that there were serious hazards to be faced in the pursuit of truth.

The second thing is that wording becomes so delicate as to be silly. Blunt comments don't happen, even when they would express a point more accurately. Besides having to dance around what is being written, the comments are now followed by lots of extra boilerplate, such as pointing out that any group can have bright people and that IQ tests are not deterministic. I must admit that I have fallen into this protective kind of language (at least when I write something that could cause blowback).

Jacobsen: What will happen to respected disciplines where international standing matters with individuals selected in such a manner?

Williams: So far, we are in a mode of having some people who are willing to take on dangerous topics and those who will not. Although there are only a few researchers who are willing to research race and sex differences, they seem to me to be doing good work. I don't think their work has actually harmed the reputations of the nonparticipants, I have seen examples of people feeling as if they were unfairly grouped with the not-woke researchers.

Jacobsen: Truly intelligent kids will use their intelligence in one way or another. What will likely happen to these smart kids without guidance and support?

Williams: A case can be made that not supporting bright students will result in them not reaching the levels of performance that would more likely be reached with support. As you observed, bright students will pursue their interests, despite barriers from school administrators and politicians. Douglas Detterman, founder of ISIR and *Intelligence*, wrote a good article pointing out that 90% of the variance in educational outcomes is due to the individual students (intelligence). The remaining variance is split between teachers and schools, with teachers accounting for 1 to 7% of the variance. This is one of those things that lots of people will want to challenge, but Detterman has the research findings on his side.

I can't imagine what the consequences will be if the present rate of irrational policies in education continue to increase. The people who are driving things, such as equal outcomes, apparently have no idea of the magnitude of the bell curve range. Yet, they are pushing to really have college educations for every child of every ability level. Economically and practically, this is insane.

Jacobsen: How are you defining woke here?

Williams: "Woke" has become the tag for the left, with all of the policies that they push (socialism and irresponsible spending on things that are waste). In the things I have been discussing, I use "woke" in reference to policies that relate to education, such as the canceling of gifted programs; the failure to recognize student achievement out of fear that a nonachiever might feel bad about his [or her] failure; school administrator embarrassment over the suggestion that a student is brilliant; etc.

Recently Thomas Jefferson High School for Science and Technology was denied the use of tests for admission. The student body has typically been about 70% Asian, 20% White, and 2% Black, with the balance consisting mostly of Hispanic. The school board ruling that they could not use tests was challenged and went through the state justice system. The school lost. Then it was appealed to the Supreme Court but was not accepted, despite their willingness to rule against Harvard for similar discrimination against Asians.

The links below are largely redundant. They report the court's choice.

https://virginiamercury.com/2024/02/20/supreme-court-wont-hear-thomas-jefferson-admissions-case/

https://reason.com/volokh/2024/02/20/supreme-court-refuses-to-hear-case-involving-use-of-race-neutral-means-to-facilitate-anti-asian-discrimination-at-selective-public-high-school/

https://www.edweek.org/policy-politics/supreme-court-declines-case-on-selective-high-school-aiming-to-boost-racial-diversity/2024/02

https://thehill.com/regulation/court-battles/4478329-supreme-court-racial-discrimination-challenge-tj-high-school-admissions/

Now the school must admit on the basis of race, not ability. They are in a bind. If they maintain their former standards, they will have to fail most of the quota students. If they are afraid to fail them (most likely), they will have to either provide an easy option for them or simply award diplomas for attending classes.

Jacobsen: An assumption: censorship of research tends to make people – of all stripes – become creative and then pursue different means by which to explore the original subject matter. Smart, creative people are forced to get more creative and use their intelligence more. With a discouragement and a reduction in focus on general intelligence and on IQ in formal

tests, how are intelligence researchers pursuing paths for measurement of intelligence if at all? I am making a historical extrapolation as if it will happen or has already happened, potentially a bias to be optimistic about researchers and intellectual pursuits. (I'm sorry!)

Williams: At the last ISIR conference, one of my friends wondered out loud if animal studies could be used to show the things that are so obvious among humans, then use the findings as comparisons to human behaviors. Curiously, we already have a very wide range of intelligence in dogs that is quite similar to the range seen in people. There are border collies at the top and Afghan wolfhounds at the bottom.

I think the twist that might not be anticipated by the anti-intelligence faction, is AI. [Mentioned previously.]

Jacobsen: What were magic multipliers? The term "magic" tells a bit of the story.

Williams: It came from this paper: Dickens, W.T. and Flynn, J.R., 2001. Heritability Estimates Versus Large Environmental Effects: The IQ Paradox Resolved. *Psychological Review*, 108(2), p.346. In the paper, Dickens and Flynn described their imagined explanation for how imagined environmental effects could cause large impacts on intelligence. Their argument was reminiscent of the "butterfly effect" which was used in the discussion of weather. With no supporting data, the authors invented a process that they claim could convert tiny unobserved environmental effects into large factors that impact intelligence. After the inane model was offered, there were no publications showing anything that could possibly support the model. I called their model "magic multipliers" because that describes their invention. To me, this is much like inventing a story where Noah builds an ark and stocks it with two of every species, so that the flood story can be supported.

Jacobsen: Why did Plomin stop giving updates every 2 years?

Williams: Probably because the SNPs were found. I don't recall that he ever spoke to ISIR after the breakthrough that he details in Robert Plomin – Blueprint: *How DNA Makes Us Who We Are*, Penguin Books Ltd., 2018, ISBN 9780241282076.

[Editors' Note: http://tankona.free.fr/plomin2018.pdf]

ISIR honored Plomin with the Lifetime Achievement Award in 2011. He spoke to ISIR in 2013 (Cypris), but I did not attend because of the very remote location. I recall (sitting a few feet away) that he received the Distinguished Career Interview, but I am not sure of the year. By 2018 the new age of genetics arrived. Besides Blueprint (above) there is a related paper that is worthwhile: Plomin R, von Stumm S. The New Genetics of Intelligence. *Nat Rev Genet*. 2018 Mar;19(3):148-159. doi: 10.1038/nrg.2017.104. Epub 2018 Jan 8. PMID: 29335645; PMCID: PMC5985927.

Jacobsen: If the FE is decoupled from *g*, as in not a JE, how much is the decoupling – complete, or is it on a sliding scale depending on context?

Williams: My take, as of today, is that the decoupling is close to total, but there are suggested FE causes that should show some *g* loading. One example would be a decrease in mean family size. If this were to happen (it obviously has happened at the high end), it should be largely due to smaller low IQ families. That would cause a real gain in intelligence, which would probably be little more than a recovery of the already lower mean due to the negative correlation between IQ and fertility rate. Besides just hitting the low end of the IQ spectrum, there is also a small birth order effect. A reduction in family size would mean fewer children born with high birth order numbers. These children are statistically less intelligent than their older siblings. I don't think either of these have been demonstrated to show a FE.

It is a bit frustrating to see the large number of references to the FE accompanied by comments that the population is becoming more intelligent. The opposite is happening. People simply do not understand that the FE is a time and location effect that can be positive or negative at any given observation; that it is not always up; and that it is rarely (or never) a Jensen Effect.

Jacobsen: Are societies giving screening of gametes for parents with reproductive issues, single parents with means who select surrogates or sperm donors based on verified characteristics, or individuals who want to know risk factors associated with their reproductive capabilities in genetics alone, making an ethical decision in conscious, evidence-based, reasoned reproduction in a non-totalitarian, democratic fashion? Is this likely to become widespread? It's, in a way, a more precise form of how individuals engage in sexual selection in the first place happening for millennia.

Williams: That takes in a lot! It is my understanding that IVF usage is large in some nations and varies down to zero in many nations. I am not familiar with the policies of the nations where IVF is most prevalent. I looked at the web and found that the US has 1.7% of all infants born through Assisted Reproductive Technology, whereas Denmark has an estimated 8% to 10% conceived through ART. That strikes me as a relatively large fraction. It seems that IVF or ART might be used more in the future, but by educated people. It is difficult for me to imagine it as equally attractive for low IQ families.

Jacobsen: Once we get the structure and networks and processes most likely connected to *g* in the brain, what would this mean for the development of simulations of this in computers, artificial *g*?

Williams: It is difficult to rule anything out for the future. The rate of development of computer technology remains high. The expected diminishing returns are being crushed by new technologies. We already see optical technology that claims to offer petabytes of storage on an optical disk that is the size of the old ones we have mostly discarded. [Using that kind of storage may be another matter, but we keep thinking of barriers that fall.] And we have been seeing research in quantum computing for some time. It seems to be real and progressing towards ultimate implementation. With what appears to be unlimited speed and storage, plus AI, getting to the point of using brain structures and processes in computers may be a matter of time.

Some time ago, I read a paper [Jung, R.E. and Haier, R.J., 2007. The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. *Behavioral and Brain Sciences*, 30(2), pp.135-154.] that discussed what the brain is doing with information that gives us the neurology of *g*. The answer, in part, is that the brain carries out an information integration process, that is either *g* or is strongly related to *g*. In 2007, there was limited understanding of networks, as compared to today. I have not seen a merging of modern network models with the Parieto-Frontal Integration Theory, but I think there are papers that attempt to update the P-FIT model.

The Older Population is More Cognitively Able Than in the Past and Age-related Deficits in Cognition are Diminishing Over Time

Stephen P. Badham

Abstract

Increased lifespan in the population was historically driven by reductions in infant mortality but is now driven by reduced mortality in older adults. Research is beginning to show reduced incidence of many age-related diseases, but there have been some mixed trends observed in assessment of cognitive ability in healthy ageing research. Across three studies, time-based trends in older adults' cognition were assessed. In a meta-analysis of literature largely studying different waves of longitudinal data, Study 1 showed cognitive improvement in later-recruited waves of older adults compared to earlier-recruited waves. In a second meta-analysis of studies comparing young and older adults' cognition, Study 2 showed that age-related cognitive deficits were becoming smaller over time. Finally, in an analysis of historic data from a single laboratory, Study 3 confirmed the findings of Study 2 and demonstrated that time-based reductions in age-related cognitive deficits were largely driven by improvement in cognition over time in older groups, whilst young adults' cognition remained relatively flat across time. It is argued that later tested groups of older adults are benefiting from environmental advantages to cognition (e.g., education, healthcare, nutrition) that might previously have mainly applied to young adult groups in research. These results have implications for cognitive ageing research which will likely yield smaller age differences than historic work. It is also argued that definitions of cognitive impairment related to dementia diagnosis may need to be periodically revised.

Introduction

Global reductions in mortality are causing a shift in population demographics, which is leading towards a greater proportion of older adults (65+) in society (United Nations, Population Division, 2022). To project the impact of this demographic shift, it is important for researchers to establish the epidemiology of ageing to effectively inform public health policy and allocation of resources (Gao et al., 2019; Jaul & Barron, 2017). This is a particularly challenging task as factors driving changes in mortality are also changing what it means to become old. For example, recent literature indicates a reduction in the incidence of several age-related health issues over time including dementia (Gao et al., 2019; Wu et al., 2017; Farina, Zhang et al., 2022), stroke (Madsen et al., 2020) and heart failure (Groenewegen et al., 2020). In contrast, other disorders have shown a rise in incidence over time such as depression over age 50 (Weinberger et al., 2018), liver cancer (Dasgupta et al., 2020; Liu et al., 2020) and hearing loss (Wallhagen et al., 1997). The current article focuses on establishing if cognitive ability is improving or declining over time in healthy older adults.

Cognitive ageing represents one of the most established fields of ageing research that systematically evaluates healthy older adults. The influence of cognition on behavior, health and wellbeing is evident in a variety of ageing research. Leading theories have

documented age deficits in cognition such as the slowing of information processing (Salthouse, 1996), a reduction in ability to inhibit or suppress information (Hasher and Zacks, 1988), and widespread deficits in ability to memorize information (Fraundorf et al., 2019; Old & Naveh-Benjamin, 2008; Rhodes, Greene, & Naveh-Benjamin, 2019). These cognitive deficits are linked to many health-related risk factors, most notably, pathological cognitive decline is a determining factor in the diagnosis of dementia (Salmon & Bondi, 2009). Additionally, cognitive deficits in healthy adults are also associated to sensory loss (Roberts & Allen, 2016), brain atrophy such as shrinkage of the prefrontal cortex and decreases in grey matter and white matter (Park et al., 2022), as well as general health and cardiovascular disease (Deary et al., 2009). In terms of individual behavior, reduced socialization (Charles & Carstensen, 2010) reduced physical activity (Erickson & Kramer, 2009) and low-quality diet (Feart et al., 2009) are all associated to lower cognitive ability in older adults.

In addition to building upon literature showing trends in ageing, there are multiple practical reasons to establish potential changes in age-related cognitive deficits over time: Firstly, ageing researchers may need to revise the expected deficits for samples of a given age when evaluating theories of ageing and cognition, when replicating prior research, and when associating cognitive deficits to other health factors. Secondly, at a personal and public interest level, a loss of cognitive ability represents a key fear of ageing reported by individuals (Brunton & Scott, 2015), which has been shown to impact quality of life in older adults (Farina, Bennett et al., 2022). Thirdly, cognition is linked to an individual's ability to live independently both in healthy older adults (Cahn-Weiner, Boyle, & Malloy, 2002) and in dementia (Martyr & Clare, 2012), providing a key indicator of the global health burden of ageing: Time-based trends related to independence show that whilst rising lifespan is associated with a greater duration of health, it is also associated with an even greater period of poor health, with evidence for an increase in years of dependency before mortality across time (Kingston et al., 2017). Finally, given the trend of lower dementia incidence over time (Gao et al., 2019; Wu et al., 2017; Farina, Zhang et al., 2022) it is crucial to establish if healthy cognitive ageing is also changing over time, so that dementia diagnosis thresholds accurately correspond to pathological cognitive decline in excess of decline due to healthy ageing.

The current article comprises three studies to investigate three different evidence bases for time-based trends in older adults' cognitive ability. In Study 1, a meta-analytic review was conducted on existing studies with data suitable for establishing time-based trends in older adults' cognitive ability. This largely included multi-wave longitudinal research where latter recruited waves could be compared to earlier recruited waves using the same cognitive measures. Many studies already report time-based trends in older adult's cognition, however, some show older adults' cognition improving in latter cohorts compared to earlier (e.g., Gerstorf et al., 2023; Dodge et al., 2014; Rönnlund & Nilsson, 2008; 2009), some show the reverse (e.g., Hülür,et al., 2013; Oi, 2017), and some show similar cognitive ability across time in older adults (e.g., Graves et al., 2021; Overton et al., 2018). Therefore, the review aimed to summarise the overall pattern across this literature. In Study 2, a meta-analysis was conducted on studies that contrasted young and older adults on cognitive ability across a much larger variety of research. This was to determine if age *differences* are changing over time, which is a novel alternative source of information on trends in age-related change. Notably, this approach also investigates the same cross-sectional paradigms that are most common in the field of cognitive ageing research and informs on cognitive differences that might be expected in ongoing investigations. Finally, Study 3 analysed the author's own historic data which included (i) information on time-based trends by using the same measures taken at different points in time and (ii) age differences between young and older adults at each time point. Together, these studies aimed to determine if older adults' cognition is changing over time, and implications of such trends for researchers and wider society.

Study 1: Cohort effects on cognition in older adults

An established pattern in the literature is that cognitive ability in the general population is improving over time, most notably in studies of IQ. This has been termed the Flynn effect after the trend was established by James Flynn in seminal large-scale studies (e.g., Flynn, 1984; Flynn, 1987), and continues to be observed in more modern research (e.g., Wongupparaj et al., 2023). The Flynn effect has also been investigated and observed in populations of older adults (e.g., Hessel et al., 2018; Skirbekk et al., 2013). A variety of research seems to show that later-born cohorts of older adults cognitively outperform earlier-born cohorts of older adults, but as highlighted in the introduction above, sometimes the Flynn effect is not present. Research in this area predominantly utilizes historic analysis of existing longitudinal databases and compares participants from different waves of recruitment at the point of first testing. This allows a cross sectional comparison of time-based trends in older adults' cognition without the confound of practice effects. As the effect is not always consistent, the current study aimed to quantitatively discern the overall pattern of research in this field. By synthesizing results across a variety of studies, the current research was also able to establish if changes in older adults' cognitive ability over time correspond to the amount of time separating comparable cohorts; this would provide additional evidence for the Flynn effect beyond averaging the results of different studies.

Method

Study selection

This study aimed to review and synthesize existing literature investigating cohort-related trends in cognition in older adults. This would establish if the Flynn effect was ubiquitous across a variety of longitudinal work, where the same cognitive measures were taken across time-separated cohorts.

An initially narrow screening search ('age AND cogni* AND cohort') was conducted on Web of Science for articles on 27th May 2023 which led to 432 results. Titles and, where relevant, abstracts for these articles were read identifying 13 articles of interest. This

was followed by a much larger snowball search of the references and citations of each of these articles for additional relevant work. As each new article was identified, the references and citations of every article were checked until no new relevant articles appeared. Articles published before the year 2000 and articles with a focus on unhealthy ageing (using dementia-focused cognitive assessment such as the MMSE) were not considered for this analysis and review (see Stephan et al., 2022 for a review of dementia-focused measures). This resulted in the identification of a total of 61 potentially relevant articles and the full text for each of these was consulted; 10 of these articles were then determined to be irrelevant (either dementia-focused or based on a single wave of recruitment) resulting in a final review of 51 articles.

Appendix Table A1 outlines the characteristics of each study, and the supplementary materials (Badham, 2023) contain all information extracted from the original articles. Table A1 shows the article citation, the underlying longitudinal cohort on which the article is based, the geographical location of the cohort, the cognitive measures assessed, an indication of whether earlier and later recruited cohorts showed similar declines over time, the years of earlier and later recruitment waves, an indicator for if the later recruited cohort was more or less cognitively able than the earlier-recruited cohort, and an indicator to highlight if education was reported to explain some of the variance between cohorts in each study. Cognitive measures varied across the studies, and individual studies often had more than one measure, popular measures were verbal fluency (n = 21, ability to list words within a category), memory (n = 35, including delayed recall, delayed recognition, immediate memory and working memory), and cognitive speed (n = 16, including the digit symbol substitution task, see Study 3 of this article for details of this task).

Data preparation

Where possible, data was extracted for the mean and standard deviation of the cognitive measures in the earliest and latest wave of each study, these were noted alongside the year of data collection (mean year if recruitment lasted longer than a year), the mean age of each cohort, the proportion of females in each cohort and the level of education in each cohort. It was also noted if the original authors concluded a cohort-related difference in cognition, a cohort-related difference in the rate of cognitive decline in follow-up longitudinal

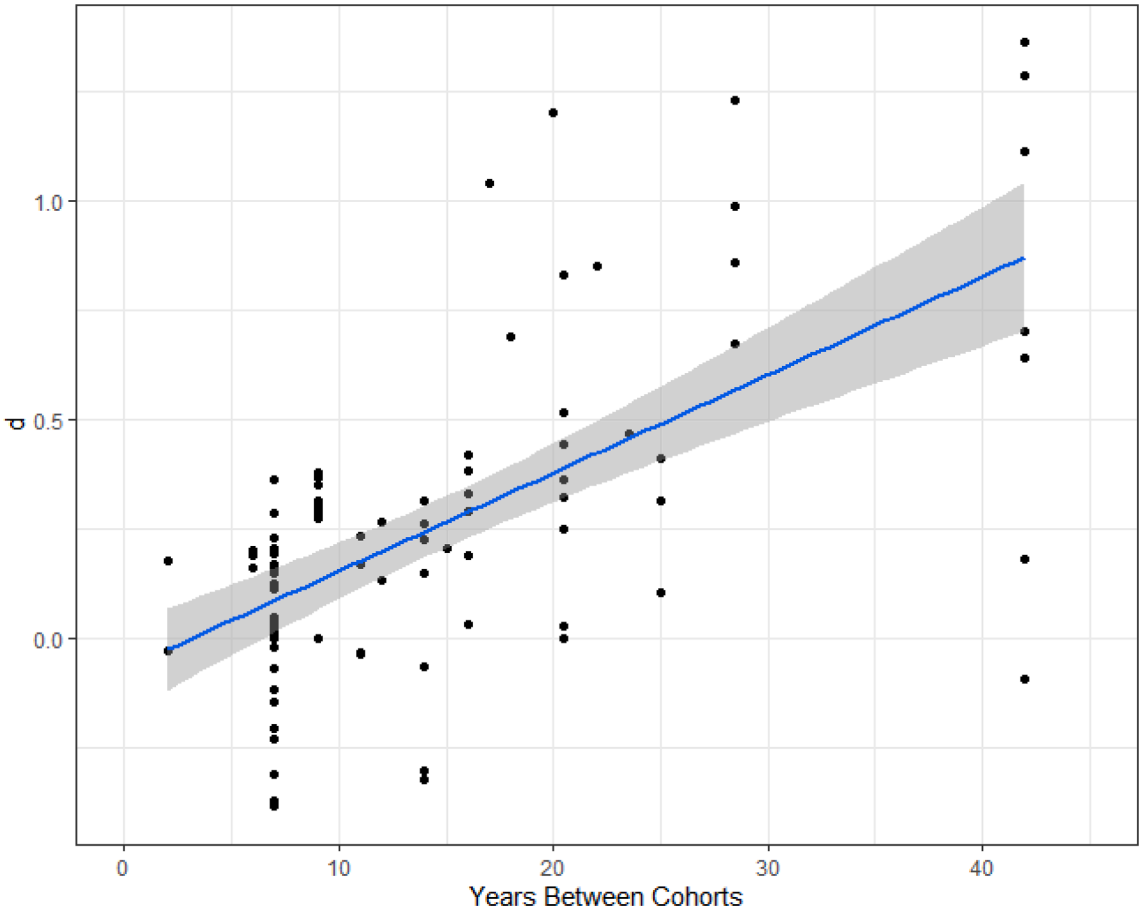


Fig. 1. Cognitive Measure Effect Size d Between Earlier-Recruited and Later-Recruited Cohorts of Older Adults in Study 1. *Note:* Positive effect sizes show cognitive advantage for later-recruited cohorts over earlier-recruited cohorts. Regression slope shows 95% CIs. Positive slope indicates greater advantage to later cohorts as time between cohort testing increases.

assessment, and the potential for education to explain the variance between cohorts.

Results

All analyses in this article are two-tailed with an alpha of 0.05 determining significance, analyses were conducted in IBM's SPSS v28. The analyses sought to establish if later-recruited cohorts show better cognition than earlier-recruited cohorts. Of the 51 articles, 45 provided a readily interpretable conclusion on cohort differences. (Six articles did not: one only reported data without statistical tests, two aimed to avoid cohort effects as a confound, one focused on sex differences only, and two focused on socioeconomic circumstances at birth.) As studies often contained multiple measures there were 123 potential cohort differences. Eighty-three measures (68 %) showed better cognitive performance in later cohorts of older adults than earlier cohorts, only six (5 %) showed the reverse, there were also 32 (26 %) non-significant cohort differences, and two measures were significant with non-linear patterns. The odds of a later cohort showing better cognition than an earlier cohort, compared to the reverse was therefore 83/6 = 13.8, and the odds of a later cohort showing better cognition than an earlier cohort compared to any other result was 83/40 = 2.1.

From the measures reported, 98 were reported such that an effect size could be computed for the cognitive difference between the earliest cohort and the latest cohort. Cohen's *d* was computed for each of these (or taken from original articles where available), a one-sample *t*-test showed a significant cognitive benefit to later cohorts over earlier cohorts (mean $d = 0.25$ [95 %CI = 0.17 to 0.32], vs. no difference 0); overall, average cognitive performance was better in later, compared to earlier cohorts, $t(97) = 6.76$, $p < .001$. Extending upon the original data from any single previous study, Fig. 1 shows the distribution of these effect sizes plotted against the number of years between testing of the earliest and latest cohorts (93 measures), the regression slope, showing a 0.02 [95 %CI = 0.017 to –0.028] increase in effect size per year of cohort time separation, was significant ($R^2 = 0.39$, $F(1, 91) = 58.89$, $p < .001$), indicating that the magnitude of the cognitive advantage to later cohorts over earlier cohorts was greater as the two cohorts were separated further in time.

Some studies reported multiple cognitive measures and may have been disproportionately represented (e.g., Overton et al., 2018 at 7-years cohort separation reported 29 measures). To account for potential variance due to any given study, the individual papers linked to each data point were entered as a random factor to group points from the same studies. For the effect size between later and earlier cohorts as the dependent variable, and paper identity as a random factor, an initial model retained a significant intercept which favored the latter cohort, $F(1, 21) = 27.04$, $p < .001$, comparable to the one-sample *t*-test above. When the amount of time separation between cohorts was entered as a covariate to the above model, it significantly predicted the effect size, $F(1, 20) = 28.94$, $p < .001$, confirming the significant regression slope above and in Fig. 1.

Education was considered as a covariate in many of the original studies. Of the 46 occasions that evaluated education, 38 (83 %) reported that education explained some of the cognitive differences observed between earlier and later cohorts. In the current investigation, education was extracted where available for earlier and later cohorts. The measures of education varied from study to study (e.g., proportion higher education or years education), to maximize data inclusion, education for the following analysis was approximated by taking the later cohort's education measure and dividing it by the earlier cohort's education measure. A correlation showed that education was related to the time difference between the two cohorts, $r(63) = 0.383$, $p = .002$, [95 %CI = 0.151 to –0.575] but it was not related to the effect size of the cognitive difference between the earlier and the later cohorts, $r(64) = 0.186$, $p = 1.37$, [95 %CI = −0.060 to 0.411]. A hierarchical regression (see Table 1) was computed, predicting the cognitive difference between the earlier and the later cohorts. At Stage 1, education was entered as a predictor, at Stage 2, the number of years between the earlier and the later recruited cohorts was entered as a predictor. Stage 1 was non-significant but there was a significant improvement to the model in Stage 2. This analysis showed that the cognitive difference between earlier and later cohorts was significantly related to the time difference between recruitment of those cohorts after accounting for a measure of education. In addition to assessing education, the difference between the proportion of females in the earlier cohort versus the later cohort was assessed: it was unrelated to cognitive differences

Table 1
Hierarchical regression predicting cognitive change between earlier recruited and later recruited cohorts in Study 1.

		Unstandardized Coefficients		Standardized Coefficients			
		B	Std. Error	Beta	*t*	*p*	R^2
Model 1	(Constant)	−0.015	0.109		−0.140	0.889	0.038 $F(1,58) = 2.27$, $p = .138$
	Education	0.132	0.088	0.194	1.505	0.138	
Model 2	(Constant)	−0.011	0.099		−0.114	0.910	0.225 $F(2,57) = 8.27$, $p < .001$
	Education	−0.051	0.094	−0.075	−0.546	0.587	$\Delta R^2 = 0.187$ *[F(1,57) = 13.78,* *p = .001 vs. Model 1]*
	Time Difference	0.021	0.006	0.510	3.712	<0.001	

between cohorts, education differences between cohorts or years between recruitment of cohorts (*r*s < 0.25), so no further analyses were conducted.

Having established that the time between cohort recruitment favours better cognition in later recruited groups, studies with relevant longitudinal data were assessed to establish if the longitudinal *rate* of age-related cognitive decline was different for earlier-recruited or latter-recruited cohorts. Here, the reported findings across the studies were tallied up and were much less conclusive. Twelve cognitive measures showed steeper decline in the earlier-recruited cohorts compared to the latter, 8 measures showed the opposite (i.e., steeper decline in the later-recruited cohorts), 11 measures showed comparable declines in time-separated cohorts and 5 measures showed a mixture of time-separated patterns of decline. Overall, this indicates that the time-based improvement in older adults' cognition established above may not correspond to a time-based change in the longitudinal rate of cognitive decline.

Discussion

The consensus across the existing literature was in support of higher cognitive ability in later-born cohorts compared to earlier-born cohorts of older adults. Additionally, the advantage of later-born cohorts over earlier born cohorts was greater when more time separated the two groups, suggesting a continuous improvement in cognition over time. This continuous function indicates that multiple factors underly the Flynn effect such as ongoing improvements in education and healthcare as opposed to one-off events/policy changes. Fig. 1 showed this as a 40-year trend that established a broad estimate of the Flynn effect across many nations with different underlying education and healthcare policies. As education is hypothesised to be a factor underlying the Flynn effect (e.g., Wongupparaj et al., 2023), this was explored as a potential mediator. Whilst latter cohorts had more education than earlier cohorts, controlling for education did not eliminate time-based increases in cognitive performance in the current analysis. However, in the underlying studies, where education was measured, in most cases (83 %) it did explain cognitive variance across cohorts. There was also no potential for the proportion of females (c.f. Bloomberg et al., 2021) across cohorts to explain patterns in the current analyses.

Additionally, the current review showed inconsistency in the longitudinal literature on conclusions about the rate of cognitive decline. A more conclusive message was reached in a recent meta-analysis on this specific topic which indicated little evidence that increased education attenuated the rate of cognitive decline (Seblova, Berggren, Lövdén, 2020). Similarly another study on the relationship between education and age-related neurophysiological change showed no evidence linking increased education to reduced brain ageing (Nyberg et al., 2021). One may expect the higher cognitive ability of later-born cohorts to reflect underlying good health which would correspond to a slower rate of cognitive decline (c.f., Lövdén et al., 2020), but this pattern was not consistently observed in the current article. Alternatively, it has been argued that rising lifespan across time could lead to a greater proportion of frail participants in later cohorts, who may be more likely to decline faster (Gerstorf et al., 2011; Karlsson et al., 2015). Overall, the current results indicate that cognition is improving over time in older adults, and that the Flynn effect applies to observations of older adults.

Study 2: Age differences in cross-sectional studies over time

Given that older adults' cognition was rising over time in Study 1, this may also have implications for assessing age-differences in cognition. Research has shown that a large proportion of cross-sectional age deficits in cognition can be explained by the Flynn effect (Agbayani & Hiscock, 2013; Dickinson & Hiscock, 2010). This is because as population IQ increases over time, the later born younger groups in studies are born in a period of higher IQ than the earlier born older groups with whom they are compared. Study 1 above showed that as we move into a more modern era, older adults too are benefiting throughout their lives from factors that boost cognition and drive the Flynn effect; these have been hypothesised to be factors such as better education, nutrition, sanitation, vaccination, and more cognitively stimulating environments (see Wongupparaj et al., 2023, for a review).

A crucial consideration in terms of the influence of the Flynn effect on cognitive ageing comparisons is that recent research has shown a reduction and reversal of the Flynn effect. Dutton, van der Linden and Lynn (2016) and Woodley of Menie et al., 2017 reviewed a range of research that indicated widespread reductions in IQ, largely in more recent years and those authors argued that there is a limit by which IQ can rise over time. Furthermore, in Wongupparaj et al.'s (2023) review they found that the Flynn effect was strongest in the mid-20th century as opposed to more recent years. Other research has also shown that the Flynn effect is more evident in countries with initially lower performance levels (Hessel,et al., 2018; Wongupparaj et al., 2015) indicating that the Flynn effect is larger when there is more to gain. Therefore, it may be the case that in more recent times, the advantage of young adults compared to older adults due to time-period influences on cognition may be changing. In the current study, a widespread investigation of cross-sectional cognitive ageing studies were investigated that compared groups of young and older adults. This was achieved via a meta–meta-analysis which was conducted on cognitive ageing meta-analyses reported since 2015.

Method

Article search

A search was conducted for meta-analyses that assessed cross-sectional age differences in cognition with groups of young and older adults. A cut-off point of 2015 was used when searching for articles to ensure potential for a wide time range in reported data. An exhaustive search was not intended, and the searching process and data compilation was conducted before any analyses were conducted. Using Web of Science, the search term [meta-an* AND cogni* AND old* AND young*] from 1st January 2015 to 1st Jan 2023

Table 2
Meta-analyses utilized to establish trends in age-related cognitive deficits over time reported in Study 2.

Meta-analysis	Focus	Number of sub studies	Measures	Original Data (Location)	Data Preparation
Badham et al. (2022)	Episodic detail reporting in autobiographical descriptions	25	Amount of episodic detail	Pooled standard effect of age difference (numerical values appear in their Fig. 7)	Converted to z-scores with positive values favoring older adults
Cervera-Crespo and González-Alvarez (2017)	Hayling Task, sentence completion with cued (automatic) and non-cued (inhibitory) words	11	Response latencies and error rates for automatic and inhibitory versions of the task (4 measures)	Hedge's *g* for age difference latency (Tables 2-5)	Within each measure, converted to z-scores with positive values favoring older adults
Erickson, Lampinen, and Moore (2016)	Eyewitness identification	18	Eyewitness identification accuracy	Odds of accurate identification: young/older (numerical data appear in their Figure 1)	Converted odds value to z-scores with positive values favoring older adults
Fraundorf et al. (2019)	Recognition memory	187*	Recognition accuracy	Difference in d' recognition accuracy between young and older adults (github link in method section)	Converted difference value to z-scores with positive values favoring older adults
Hayes et al. (2020)	Ability to recognize emotional expressions	81	All emotional recognition (data were not split by emotion type and presentation manipulations, there were 18+ categories with largely comparable age differences)	Hedge's *g* for age difference in emotional recognition (Supplementary Materials)	Converted to z-scores with positive values favoring older adults
Rey-Mermet, and Gade (2018)	Inhibition ability	121	Color Word Stroop Task, Color Stroop Task, Flanker Task, Simon Task, Global Inhibition, Local Inhibition, Positive Compatibility, Negative Compatibility, N-2 Repetition Cost, Stop Signal Task, Go/No-Go Tasks	Response time means for baseline and interference conditions for each age group (Supplementary Materials)	Baseline vs. interference 'costs' were computed for each age group (interference-baseline RT), then age difference in 'costs' were computed (the original meta-analysis also considers all dependent variables together in one analysis). Following this, z scores for each measure were computed with positive values favoring older adults. Also, Some articles were 'in press' in the original meta-analysis, the final publication dates of those articles were obtained.
Rhodes, Greene, and Naveh-Benjamin (2019)	Recognition and recall memory	36	Recall and recognition accuracy in young and older adults	Hedges g for age differences (Appendix Table 2)	Within each measure, converted to z-scores with positive values favoring older adults
Theisen, et al. (2021)	Drift rate (speed) of data accumulation	21	Perception, Lexical Decision, Memory	Hedges g for age difference in drift rate (Supplementary Figure 1; numbers from datapoints in figure obtained from the author)	Within each measure, converted to z-scores with positive values favoring older adults
Vallesi et al. (2021)	Sustained attention, simple responses with a low probability alternative response	12	Sustained attention response times for young and older adults	Hedges g for age differences (numerical values appear in Figure 2)	Converted to z-scores with positive values favoring older adults

* Fraundorf et al. (2019) report 188 in their methods but the current author could only identify 187.

yielded 330 results, capturing all meta-analyses the author was aware of at the time of search (Badham et al., 2022; Bopp & Verhaeghen, 2020; Rhodes, Greene & Naveh-Benjamin, 2019). Analysis of titles and abstracts revealed 20 articles of interest.

Screening

The meta-analytical papers required a method of linking age differences to specific points in time, namely the publication dates of articles included in each meta-analysis. Of the 20 articles initially identified, 11 were excluded for the following reasons: Six focused on comparisons between groups of studies involving young and/or older adults as opposed to age differences within individual studies that could be linked to specific years (Chen et al., 2015; Chen et al., 2020; Huo et al., 2021; Oberste et al., 2021; Wang et al., 2016; Ware, Kirkovski, & Lum, 2020). Two articles focused on a comparison between conditions rather than a direct measure of performance: sleep vs waking memory (Gui et al., 2017); before and after cochlear implantation (Schafer et al., 2021). One article focused on longitudinal research (Stephan et al., 2021). One article contained a suitable data set but no means of linking individual data points to specific papers (Bopp & Verhaeghen, 2020). And one article had been retracted (Armstrong et al., 2021). The remaining nine articles reported data suitable for establishing age differences in cognition at given time points, as outlined below.

Data compilation and summary of measures

Data were extracted for the nine meta-analyses as outlined in Table 2 below. The age differences from individual papers and experiments reported in each meta-analysis were standardized using z-scores based on the type of measure (e.g., recognition accuracy, Stroop-task response time). Each z-score was associated to the year of publication of the original study (n = 512, range 1962–2022), and as studies often contained multiple experiments/measures, there were 1832 unique data points. Data were expressed such that higher z-score values represented age differences favoring older adults.

Where the same measure appeared in two different meta-analyses, it was checked if the measures were the same dependent variable. If they were, the data point from the smaller meta-analysis was deleted (after computing the z-scores, which remained based upon the undeleted means and SDs of effect sizes for that measure), to preserve the study effect best in the context of the literature. Despite two meta-analyses being based off recognition, only 14 age differences were removed. The extracted data can be found in the supplementary materials (Badham, 2023).

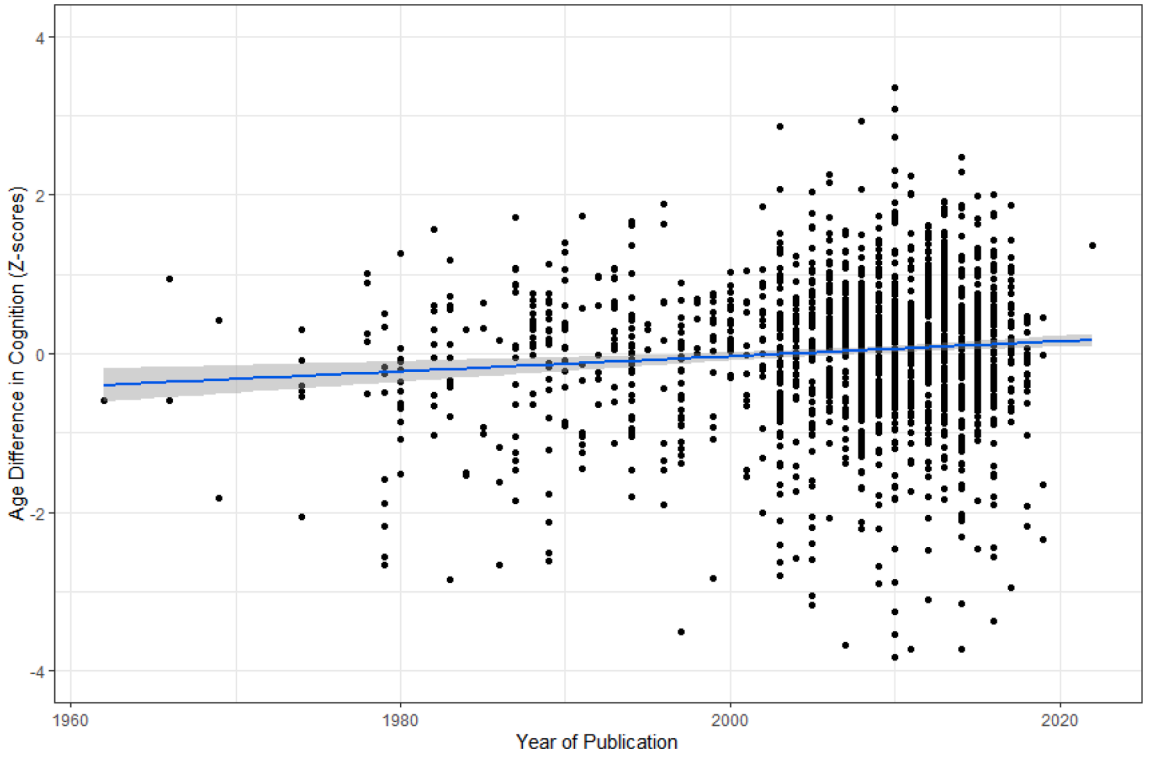


Fig. 2. Z-score of Age Difference in Cognition Plotted Against Year of Publication from Study 2. *Note:* Regression slope shows 95% CIs. Positive slope indicates reduced age deficits with increasing time. The regression slope of 0.009 corresponds to a 0.54 Z-score change across the 60-year range such that that the average age deficit in 2022 would be approximately in the bottom 30% of age deficits from 1962. Three outlying data points are excluded from image only for clarity (2010, −10; 2010, −10; 2010, −7).

Results

For the primary analysis, there was a significant correlation between year of publication and magnitude of age deficit such that age deficits diminished over time, $r(1830) = 0.081$, $p < .001$, [95 %CI = 0.036 to 0.127], (see Fig. 2). To account for possible duplicate samples or biases arising from individual studies with many measures, this analysis was repeated with just one data point from each of the 512 papers and the result held, $r(500) = 0.102$, $p = .021$, [95 %CI = 0.015 to –0.187] (this analysis was only conducted once following one arbitrary selection of samples). As the meta-analyses each featured different ranges of years of publication, a mixed effect model allowing different intercepts for each measure, but one fixed slope, was applied to the data. This also showed a significant relationship between year of publication and age deficits In the same direction as above: with the full data, slope = 0.009, $t(1830) = 3.49$, $p < .001$, [95 %CI = 0.004 to –0.014], indicating a reduction of age deficits of 0.009 SD per year; and with a data point form each paper, slope = 0.011, t(273.50) = 2.36, p = .019, [95 %CI = 0.002 to –0.020].

As it has been hypothesized that changing health trends and ability over time could be driven by life expectancy (Gao et al., 2019; Kingston et al., 2017), data from the United Nations Population Division was used to establish the life expectancy at age 65 for each year of publication. The life expectancy data for USA was used as it represented the location of most of the research and was comparable to other countries conducting research (patterns of significance were the same for world life expectancy). The life expectancy numbers correlated with the magnitude of the age deficit in cognition, such that age deficits were smaller when life expectancy was longer, $r(1830) = 0.079$, $p < .001$, [95 % CI = 0.033 to –0.124], (see Fig. 3). Life expectancy may have moderated the influence of year of publication on age deficits, but life expectancy and year of publication were too highly correlated for that analysis, $r(1830) = 0.981$, $p < .001$, [95 % CI = 0.980 to –0.983].

Discussion

This study showed that time-based trends in cognition are likely working differently for young and older adults, given that across a large variety of studies age-related deficits in cognition are diminishing over time. This is particularly important on a theoretical basis for cognitive ageing research as it means that effects detectable in the past may disappear or require much larger sample sizes to detect in ongoing research. It also means that the decline an individual might expect to experience as they become older is smaller than originally thought, which is consistent with studies showing smaller age deficits in longitudinal compared to cross sectional research (e.g., Salthouse, 2009; Schaie, 2012). The observed trend of diminishing age deficits over time was also significantly related to a rise in lifespan which is indicative that it is driven by time-based changes in older adults more so than time-based changes in young adults, as

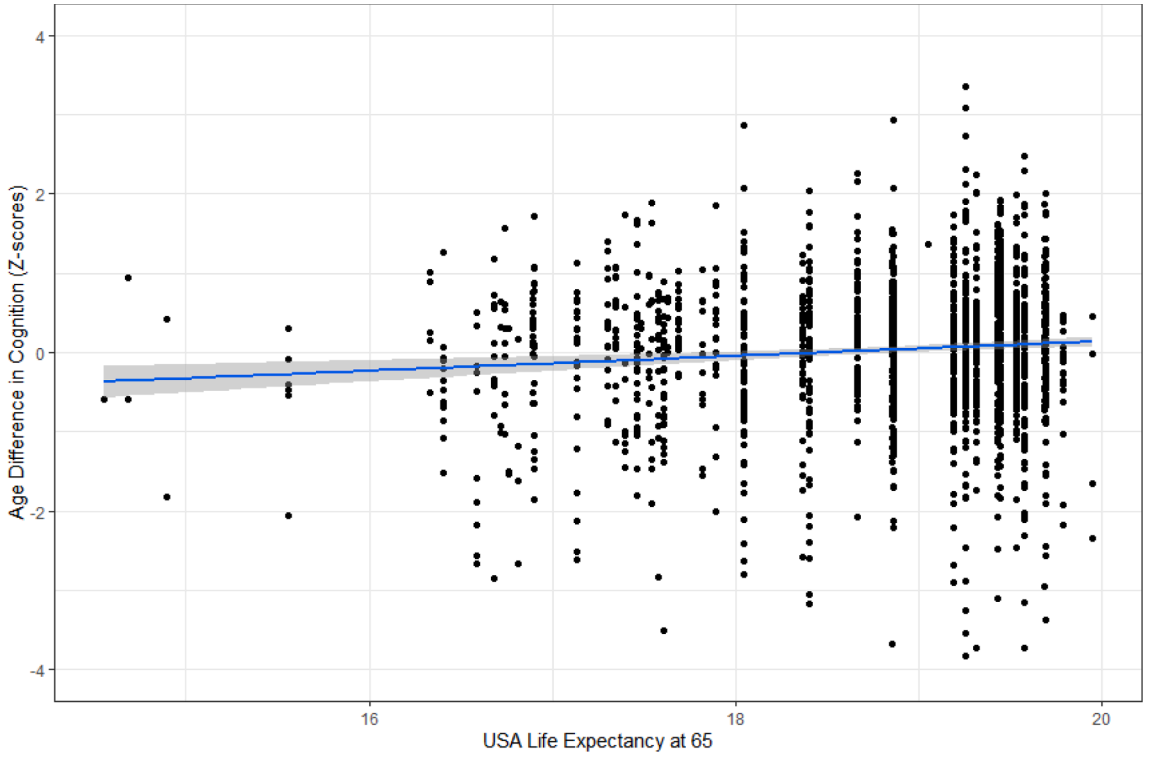


Fig. 3. *Z*-score of Age Difference in Cognition Plotted Against USA Life Expectancy at Age 65 from Study 2. Note: Regression slope shows 95% CIs. Positive slope indicates reduced age deficits with increasing life expectancy. Three outlying data points are excluded from image only for clarity (19.3, −10; 19.3, −10; 19.3, −7).

Table 3
Correlations (r) between year of data collection (2006–2009) and measures taken for each age group in Study 3.

	Age Group	
Measure	Young (N)	Older (N)
DSST (cognitive speed)	0.058 [95 %CI = −0.026 to 0.141] (543)	0.219[2] [95 %CI = 0.138–0.297] (546)
MHA (vocabulary)	0.123* [95 %CI = 0.040–0.205] (543)	0.243[2] [95 %CI = 0.162–0.320] (546)
Sex	−0.012 [95 %CI = −0.096 to 0.072] (543)	−0.012 [95 %CI = −0.095 to 0.072] (546)
Health	−0.040 [95 %CI = −0.156 to 0.077] (283)	−0.015 [95 %CI = −0.133 to 0.102] (279)
Education	0.046 [95 %CI = −0.044 to 0.136] (476)	0.271[2] [95 %CI = 0.183–0.354] (453)
Age	0.033 [95 %CI = −0.053 to 0.116] (543)	−0.008 [95 %CI = −0.092 to 0.076] (546)

* $p < .01$.
[2] $p < .001$.

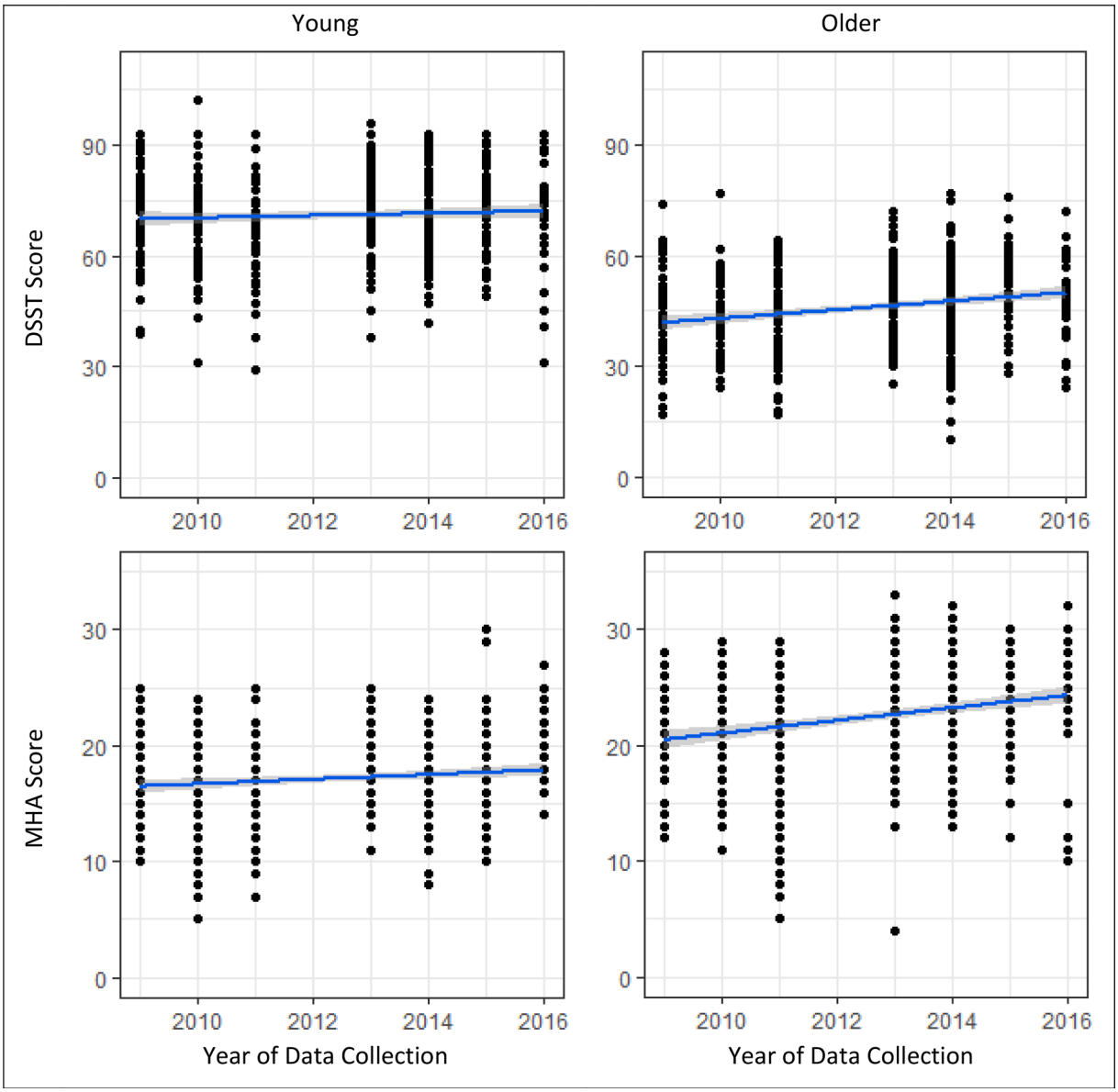


Fig. 4. Digit Symbol Substitution Task (DSST) Speed Scores and Mill Hill A (MHA) Vocabulary Scores for Young and Older Adults from Years 2009 to 2016 in Study 3. *Note:* Regression slopes show 95 % CIs. DSST Score represents number of symbols copied in 90 s and higher values correspond to faster performance. MHA Score is number of correct responses and higher values represent better vocabulary scores.

Table 4
Regression output predicting cognitive speed (DSST) in Study 3 (N = 561).

	Unstandardized Coefficients		Standardized Coefficients			
Predictor Variable	*B*	Std. Error	Beta	*t*	*p*	R^2
(Constant)	44.255	3.906		11.329	<0.001	0.597F (6, 554) = 136.97, p < .001
Sex	−4.229	0.921	−0.125	−4.594	<0.001	
Health	0.391	0.715	0.016	0.547	0.585	
Education	0.933	0.151	0.168	6.186	<0.001	
Age Group	−27.589	2.039	−0.844	−13.527	<0.001	
Year of Data Collection	−1.206	0.615	−0.053	−1.960	0.050	
Age Group × Year of Data Collection	−2.825	1.228	−0.141	−2.301	0.022	

Table 5
Regression output predicting vocabulary ability (MHA) in Study 3 (N = 561).

	Unstandardized Coefficients		Standardized Coefficients			
Predictor Variable	*B*	Std. Error	Beta	*t*	*p*	R^2
(Constant)	13.119	1.385		9.472	<0.001	0.45F (6, 554) = 74.57, *p* < .001
Sex	0.899	0.326	0.088	2.753	0.006	
Health	0.614	0.253	0.084	2.422	0.016	
Education	0.325	0.053	0.194	6.084	<0.001	
Age Group	6.294	0.723	0.637	8.704	<0.001	
Year of Data Collection	0.316	0.218	0.046	1.450	0.148	
Age Group × Year of Data Collection	−0.171	0.435	−0.028	−0.393	0.695	

modern changes in lifespan are driven by lifespan increases in the latter period of life (Mathers et al., 2015). It is also possible that the diminishing cognitive deficits observed were driven by changes in gerontological science, with researchers testing more subtle effects in later years as theories developed in complexity. Study 3 aimed to avoid complications derived from the heterogeneity of research in Study 2 by utilizing age differences measured in the same way from the same laboratory over time.

Study 3: Age differences in the same laboratory over seven years

Given the aforementioned evidence for a negative Flynn effect in modern society with some populations of young adults showing declining IQ (e.g., Dutton, van der Linden & Lynn, 2016), it may be possible that the observed effect on cross sectional work was driven by a change in young adults' ability rather than due to an improvement in older adults' ability. Study 3 aimed to simultaneously explore age differences in cognition alongside time-based trends in cognition to investigate this possibility further.

Method

Data preparation

The current author has collected similar background measures in cross sectional ageing research over a period of years, so from these data it is possible to assess changes in cross sectional age differences over time, where all measures were identical. Participants' scores for the Digit Symbol Substitution task (DSST, Wechsler, 1981) and the Mill Hill vocabulary test part A (MHA, Raven et al., 1988) were extracted from all available relevant data possessed by the author. The DSST is a measure of cognitive speed based on copying symbols as swiftly as possible from a key into a grid by hand. The MHA is a synonym-based multiple-choice measure of vocabulary. It is worth highlighting that vocabulary is a task where older adults typically show an advantage relative to young adults (Verhaeghen, 2003) and the current data are no exception. Therefore, the current analyses assess time-based changes in the age-related *advantage* in cognition in this instance.

The data was collected between 2009 and 2016 face to face in the UK (9 articles: Badham & Maylor 2011; Badham & Maylor, 2012; Badham, Estes & Maylor, 2012; Badham et al., 2013; Badham & Maylor, 2016; Badham et al., 2016; Badham et al., 2017; Badham, Sanborn, & Maylor, 2017). Information on ethnicity was not reported in the original articles but the samples were predominantly white-British. Most of the experiments within these articles compared around 30 young and 30 older adults, there were 15 experiments (with distinct data sets) in total and 1089 participants. The mean age of the young group (N = 543, 342 Female) was 20.74 (*SD* = 2.38) and the mean age of the older group (N = 546, 314 Female) was 73.29 (*SD* = 6.33). Data on participant's sex, health (self-rated on 5-point scale, higher scores show better health), and years of education were extracted, where available, all extracted data can be found in the supplementary materials (Badham, 2023).

Table A1
Summary of articles from Study 1.

Article	Longitudinal Cohort	Cohort Location	Cognitive Measures	Time by Cohort Interaction	ER* Year	LR* Year	Later better**	Education Covariate***
Ahrenfeldt et al., 2018	Survey on Health, Ageing and Retirement in Europe	Europe	Verbal Fluency + Free recall (Composite)		2004	2013	1	1
Alwin & McCammon, 2001	General Social Survey	USA	Verbal Synonym task		1972	1998		
Bancks et al., 2019	National Health and Nutrition Examination Survey	USA	Digit Symbol Substitution Test (DSST)		1999	2014	1	1
Beller et al., 2022	German Ageing Survey (Deutscher Alterssurvey; DEAS)	Germany	DSST		2002	2014	1	1
Ben-David et al., 2015	Their own Lab	Canada	Vocabulary		1995	2010	−1	1
Bloomberg et al., 2021	English Longitudinal Study of Ageing	UK	Verbal Fluency		2002	2014		
	Whitehall II study	UK	Verbal Fluency		1997	2015		
Bordone et al., 2015	German Socioeconomic Panel	Germany	Symbol–Digit Test (similar to DSST; STD90 version)		2006	2012	1	0
	English Longitudinal Study of Ageing	UK	Animal-Naming Task (Verbal Fluency)		2002	2008	1	0
	English Longitudinal Study of Ageing	UK	Immediate Recall		2002	2008	1	0
	English Longitudinal Study of Ageing	UK	Delayed Recall		2002	2008	1	0
	English Longitudinal Study of Ageing	UK	Letter-Cancellation-Task (speed)		2002	2008	1	0
Brailean et al., 2018	Longitudinal Ageing Study Amsterdam	Netherlands	Immediate Free Recall	Mixed results showing some steeper and shallower decline across time across cohorts	1992.5	2003.5	ns	1
	Longitudinal Ageing Study Amsterdam	Netherlands	Delayed Free Recall	Mixed results showing some steeper and shallower decline across time across cohorts	1992.5	2003.5	ns	1
	Longitudinal Ageing Study Amsterdam	Netherlands	DSST	Steeper decline in later recruited cohort	1992.5	2003.5	1	1
	Longitudinal Ageing Study Amsterdam	Netherlands	Inductive Reasoning (Raven Colored Progressive Matrices)	Same for both cohorts	1992.5	2003.5	1	1
Choi et al., 2018	Health and Retirement Study	USA	Composite Cognitive Score		1998	2014	ns	
Christensen et al., 2013	Their own Lab	Denmark	Composite Cognitive Score		1998	2010	1	Implied
Degen et al., 2022	Interdisciplinary Longitudinal Study of Adult Development and Ageing (ILSE)	Germany	Free Recall		1994.5	2015	1	0
	Interdisciplinary Longitudinal Study of Adult Development and Ageing (ILSE)	Germany	Delayed Recognition		1994.5	2015	1	1
	Interdisciplinary Longitudinal Study of Adult Development and Ageing (ILSE)	Germany	Digit Symbol Test		1994.5	2015	1	1
	Interdisciplinary Longitudinal Study of Adult Development and Ageing (ILSE)	Germany	Mosaic Test		1994.5	2015	1	1
	Interdisciplinary Longitudinal Study of Adult Development and Ageing (ILSE)	Germany	Finding Similarities		1994.5	2015	1	1
	Interdisciplinary Longitudinal Study of Adult Development and Ageing (ILSE)	Germany	Visual Thinking		1994.5	2015	1	1

(*continued on next page*)

Table A1 (*continued*)

Article	Longitudinal Cohort	Cohort Location	Cognitive Measures	Time by Cohort Interaction	ER* Year	LR* Year	Later better**	Education Covariate***
	Interdisciplinary Longitudinal Study of Adult Development and Ageing (ILSE)	Germany	Concentration		1994.5	2015	ns	1
	Interdisciplinary Longitudinal Study of Adult Development and Ageing (ILSE)	Germany	Word Fluency		1994.5	2015	ns	1
de Rotrou et al., 2013	Memory clinic of Broca hospital	France	Composite Cognitive Score (Cognitive Efficiency Profile)		1991	2008	1	
Doblhammer et al., 2013	Survey of Health, Ageing and Retirement in Europe (SHARE)	Europe	Cognitive Battery		2004.5	2011		
Dodge et al., 2014	Monongahela Valley Independent Elders Study (aka, MoVIES) and the Monongahela-Youghiogheny Healthy Ageing Team study (aka, MYHAT)	USA	Cognitive Battery	Steeper decline in earlier recruited cohort	1988	2006	1	0
Dodge et al., 2017	Monongahela Valley Independent Elders Study (aka, MoVIES) and the Monongahela-Youghiogheny Healthy Ageing Team study (aka, MYHAT)	USA	Various Memory Measures	Steeper decline in earlier recruited cohort	1988	2007	1	0
Finkel et al., 2007	Swedish Adoption/Twin Study of Ageing	Sweden	Verbal, Spatial, Memory	Same for both cohorts			1	
	Swedish Adoption/Twin Study of Ageing	Sweden	Processing Speed	Same for both cohorts			ns	
Fox and Mitchum, 2014	Long Beach Longitudinal Dataset	USA	Abstraction Scores		1978	1994	1	
Frith and Loprinzi, 2019	National Health and Nutrition Examination Survey	USA	Episodic Memory (CERAD-D)		2011.5	2013.5	1	
	National Health and Nutrition Examination Survey	USA	Animal Fluency		2011.5	2013.5	ns	
	National Health and Nutrition Examination Survey	USA	DSST		1999.5	2013.5	1	
Fritze et al., 2014	Survey of Health, Ageing and Retirement in Europe (SHARE)	Europe	Composite Cognitive Score		2004	2008.5		
Gerstorf et al., 2015	Berlin Ageing Study & Berlin Ageing Study II	Germany	Digit Symbol Test		1991.5	2013.5	1	
Gerstorf et al., 2011	Seattle Longitudinal Study	USA	Cognitive Battery	Mixed: Steeper decline in earlier cohort, but steeper terminal decline in later cohort.	1956	2005	1	1
Gerstorf et al., 2023	Berlin Ageing Study & Berlin Ageing Study II	Germany	Digit Symbol Test	Same for both cohorts	1990	2010	1	1
Geyer et al., 2023	Survey on Health, Ageing and Retirement in Europe	German Samples	Verbal Fluency		2004	2013	1	1
	Survey on Health, Ageing and Retirement in Europe	German Samples	Delayed Recall		2004	2013	1	1
Graves et al., 2021	California Verbal Learning Test (CVLT-II and CVLT3)	USA	Immediate Free Recall		1999	2016.5	ns	
	California Verbal Learning Test (CVLT-II and CVLT3)	USA	Delayed Free Recall		1999	2016.5	ns	
Grasset et al., 2018	e French Personnes Agées Quid cohort	France	Isaacs Set Test (verbal fluency)	Steeper decline in earlier recruited cohort	1988	2013	1	1

(*continued on next page*)

Table A1 (*continued*)

Article	Longitudinal Cohort	Cohort Location	Cognitive Measures	Time by Cohort Interaction	ER* Year	LR* Year	Later better**	Education Covariate***
	e French Personnes Agées Quid cohort	France	Benton Visual Retention Test (visual working memory)	Steeper decline in earlier recruited cohort	1988	2013	1	1
	e French Personnes Agées Quid cohort	France	DSST	Same for both cohorts	1988	2013	1	1
Grasshoff et al., 2021	Survey of Health, Ageing and Retirement in Europe (SHARE)	Germany	Verbal Fluency		2004	2013	1	
	Survey of Health, Ageing and Retirement in Europe (SHARE)	Germany	Delayed Recall		2004	2013	1	
	Survey of Health, Ageing and Retirement in Europe (SHARE)	Spain	Verbal Fluency		2004	2013	1	
	Survey of Health, Ageing and Retirement in Europe (SHARE)	Spain	Delayed Recall		2004	2013	1	
	Survey of Health, Ageing and Retirement in Europe (SHARE)	Sweden	Verbal Fluency		2004	2013	ns	
	Survey of Health, Ageing and Retirement in Europe (SHARE)	Sweden	Delayed Recall		2004	2013	1	
Henchoz et al., 2020	Lausanne cohort	Switzerland	Trail Making Test (B-A: time difference between parts A and B)		2005	2015	ns	
	Lausanne cohort	Switzerland	Verbal Fluency		2005	2015	−1	
	Lausanne cohort	Switzerland	Clock Drawing		2005	2015	−1	
Hessel et al., 2018	Survey of Health, Ageing and Retirement in Europe (SHARE)	Europe	Immediate Free Recall		2004.5	2013	1	1
Hülür et al., 2013	Asset and Health Dynamics among the Oldest Old (AHEAD)	USA	Episodic memory composite	Steeper decline in later recruited cohort	1996	2005	−1	
Hülür et al., 2019	Seattle Longitudinal Study	USA	Spatial orientation	ns	Pre-WWII Birth	Post-WWII birth	1	
	Seattle Longitudinal Study	USA	Inductive reasoning	Steeper decline in earlier recruited cohort	Pre-WWII Birth	Post-WWII birth	1	
	Seattle Longitudinal Study	USA	Word Fluency	Steeper decline in earlier recruited cohort	Pre-WWII Birth	Post-WWII birth	1	
	Seattle Longitudinal Study	USA	Number ability	ns	Pre-WWII Birth	Post-WWII birth	−1	
	Seattle Longitudinal Study	USA	Verbal meaning	Steeper decline in earlier recruited cohort	Pre-WWII Birth	Post-WWII birth	1	
Johnsen et al., 2021	Tromsø Study	Norway	Immediate Free Recall		2001	2015.5	1	1
	Tromsø Study	Norway	Delayed Recognition		2001	2015.5	1	1
	Tromsø Study	Norway	Digit Symbol Coding Test		2001	2015.5	1	1
Karlsson et al., 2015	Gerontological and Geriatric Population Studies in Gothenburg, Sweden (H70)	Sweden	Figure Logic test	Steeper decline in later recruited cohort	1971.5	2000	1	1
	Gerontological and Geriatric Population Studies in Gothenburg, Sweden (H70)	Sweden	Block Design test	Steeper decline in later recruited cohort	1971.5	2000	1	1

(*continued on next page*)

Table A1 (*continued*)

Article	Longitudinal Cohort	Cohort Location	Cognitive Measures	Time by Cohort Interaction	ER* Year	LR* Year	Later better**	Education Covariate***
Lewis and Zelinski, 2010	Long Beach Longitudinal Study	USA	Cognitive Battery		1978	1994		
Llewellyn et al., 2009	MRC Cognitive Function and Ageing Study (1991) and English Longitudinal Study of Ageing	UK	Verbal Fluency	Similar age-related decline concluded from figure but no null evaluation tests on slopes	1991	2002	1	1
Luo and Warren, 2023	General Social Survey	USA	Vocabulary Composite (WORDSUM)		1978	2018	1	1
Munukka et al., 2021	Evergreen project, Evergreen II	Finland	Digit Span		1989.5	2017.5	1	1
	Evergreen project, Evergreen II	Finland	DSST		1989.5	2017.5	1	1
	Evergreen project, Evergreen II	Finland	Verbal Fluency		1989.5	2017.5	1	1
	Evergreen project, Evergreen II	Finland	Response Time		1989.5	2017.5	1	1
Nielsen et al., 2022	Survey of Health, Ageing and Retirement in Europe (SHARE)	Europe	Cognitve Composite		2004.5	2015	1	
Oi, 2017	Health and Retirement Study	USA	Cognitive Composite	Steeper decline in later recruited cohort	1992	2004	−1	
Overton et al., 2018	Good Ageing in Skåne	Sweden	Free Recall		2002	2009	1	
	Good Ageing in Skåne	Sweden	Free Recall		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Recognition		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Recognition		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Digit Cancellation		2002	2009	1	
	Good Ageing in Skåne	Sweden	Digit Cancellation		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Pattern comparison		2002	2009	1	
	Good Ageing in Skåne	Sweden	Pattern comparison		2002	2009	1	
	Good Ageing in Skåne	Sweden	Attention (TMT-A)		2002	2009	1	
	Good Ageing in Skåne	Sweden	Attention (TMT-A)		2002	2009	1	
	Good Ageing in Skåne	Sweden	Executive Functioning (TMT-B)		2002	2009	1	
	Good Ageing in Skåne	Sweden	Executive Functioning (TMT-B)		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Vocabulary		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Vocabulary		2002	2009	ns	
	Good Ageing in Skåne	Sweden	General Knowledge		2002	2009	ns	
	Good Ageing in Skåne	Sweden	General Knowledge		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Word Fluency (categories)		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Word Fluency (categories)		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Word Fluency (letters)		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Word Fluency (letters)		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Digit Span Forwards		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Digit Span Forwards		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Digit Span Backwards		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Digit Span Backwards		2002	2009	ns	
	Good Ageing in Skåne	Sweden	Metacognition (general knowledge confidence accuracy)		2002	2009	ns	

(*continued on next page*)

Table A1 (*continued*)

Article	Longitudinal Cohort	Cohort Location	Cognitive Measures	Time by Cohort Interaction	ER* Year	LR* Year	Later better**	Education Covariate***
	Good Ageing in Skåne	Sweden	Metacognition (general knowledge confidence accuracy)		2002	2009	ns	
Rodgers et al., 2003	Asset and Health Dynamics Among the Oldest Old Study (1993 wave here) and the Health and Retirement Study (2000 wave here)	USA	Immediate Recall		1993	2000	1	1
	Asset and Health Dynamics Among the Oldest Old Study (1993 wave here) and the Health and Retirement Study (2000 wave here)	USA	Delayed Recall		1993	2000	1	1
	Asset and Health Dynamics Among the Oldest Old Study (1993 wave here) and the Health and Retirement Study (2000 wave here)	USA	Serial 7 s		1993	2000	ns	1
Schaie, 2012	Seattle Longitudinal Study	USA	Verbal meaning	Generally steeper decline in earlier recruited cohorts across measures	1956	1998		
Schaie, 2012	Seattle Longitudinal Study	USA	Spatial orientation		1956	1998		
Schaie, 2012	Seattle Longitudinal Study	USA	Inductive Reasoning		1956	1998		
Schaie, 2012	Seattle Longitudinal Study	USA	Number ability		1956	1998		
Schaie, 2012	Seattle Longitudinal Study	USA	Word Fluency		1956	1998		
Schaie, 2012	Seattle Longitudinal Study	USA	Index of intellectual ability		1956	1998		
Schaie, 2012	Seattle Longitudinal Study	USA	Index of educational Aptitude		1956	1998		
Schaie, 2012	Seattle Longitudinal Study	USA	Inductive reasoning		1984	1998		
Schaie, 2012	Seattle Longitudinal Study	USA	Spatial Orientation		1984	1998		
Schaie, 2012	Seattle Longitudinal Study	USA	Perceptual Speed		1984	1998		
Schaie, 2012	Seattle Longitudinal Study	USA	Numeric Ability		1984	1998		
Skirbekk et al., 2013	English Longitudinal Study of Ageing	UK	Immediate Recall, Delayed Recall, Verbal Fluency		2002	2008	1	
Steiber, 2015	e German Socio-Economic Panel	Germany	Symbol-Digit Test		2006	2012	1	
Tampubolon, 2015	English Longitudinal Study of Ageing	UK	Episodic Memory	Qualitatively similar rate of decline (Fig. 3)	2002	2013	1	
Thorvaldsson et al., 2017	Gerontological and Geriatric Populations Studies in Gothenburg (H70)	Sweden	Spatial Ability	Steeper decline in later recruited cohort	1971.5	2000	1	
	Gerontological and Geriatric Populations Studies in Gothenburg (H70)	Sweden	Reasoning	Steeper decline in later recruited cohort	1971.5	2000	1	
	Gerontological and Geriatric Populations Studies in Gothenburg (H70)	Sweden	Perceptual- and motor-speed	Steeper decline in later recruited cohort	1971.5	2000	1	
	Gerontological and Geriatric Populations Studies in Gothenburg (H70)	Sweden	Picture Recognition Memory	Same for both cohorts	1976.5	2000	1	
	Gerontological and Geriatric Populations Studies in Gothenburg (H70)	Sweden	Verbal Ability	Same for both cohorts	1971.5	2000	1	
Turcotte et al., 2022	Alzheimer's Disease Neuroimaging Initiative	USA and Canada	Verbal Episodic Memory		2003		1	
	Alzheimer's Disease Neuroimaging Initiative	USA and Canada	Language and Semantic Memory		2003		1	
	Alzheimer's Disease Neuroimaging Initiative	USA and Canada	Attention Capacities		2003		1	

(*continued on next page*)

Table A1 (*continued*)

Article	Longitudinal Cohort	Cohort Location	Cognitive Measures	Time by Cohort Interaction	ER* Year	LR* Year	Later better**	Education Covariate***
	Alzheimer's Disease Neuroimaging Initiative	USA and Canada	Executive Functions		2003		1	
Vonk et al., 2019	Washington Heights Inwood Columbia Ageing Project	USA	Memory Composite	Steeper decline in earlier recruited cohort			1	
	Washington Heights Inwood Columbia Ageing Project	USA	Visio-Spatial Composite	Steeper decline in earlier recruited cohort			1	
	Washington Heights Inwood Columbia Ageing Project	USA	Language Composite	Steeper decline in earlier recruited cohort			1	
Yang and Land, 2006	General Social Survey	USA	Vocabulary Composite (WORDSUM)		1974	2000	Bimodal peaks across birth years	
Zelinski and Kennison, 2007	Long Beach Longitudinal Study	USA	Reasoning		1978	1994	1	
	Long Beach Longitudinal Study	USA	List Recall	Steeper decline in later group for old-old only	1978	1994	1	
	Long Beach Longitudinal Study	USA	Text Recall	Steeper decline in later group for old-old only	1978	1994	1	
	Long Beach Longitudinal Study	USA	Space (Figure and Object Rotation)		1978	1994	1	
	Long Beach Longitudinal Study	USA	Vocabulary	Steeper decline in earlier recruited cohort	1978	1994	ns	
Zheng, 2021	Health and Retirement Study	USA	Cognitive Composite		1996	2014	Mixed: inverted u-shape with mid cohort best	
Zhou, 2022	Survey of Health, Ageing and Retirement in Europe (SHARE)	Europe	Verbal Fluency		2006	2015	1	1

* ER/LR = Earlier Recruited/Later Recruited Cohorts.
** Colum indicating if LR shows better cognition than ER, 1 = LR > ER, −1 = ER > LR, ns = no sig difference between ER and LR, blank not measured/reported.
*** Column indicating if education explained some variance in cognition between ER and LR, 1 = yes, 0 = no, blank not measured/reported.

Results

Table 3 shows correlations of cognition and demographics against year of publication, separately for young and older adults. Positive correlations revealed that the older adults showed a significant improvement over time for Speed (DSST) and Vocabulary (MHA) and the young adults showed a significant improvement over time for vocabulary. Later-tested older adults were also more likely to have more years of education. Fig. 4 shows linear regression slopes for year of data collection against each cognitive measure for each age group.

To establish if changes in cognitive scores over time were different for young and older adults, and if this could be accounted for by gender, health and education, two regressions were conducted (one for DSST as an outcome variable, see Table 4, and one for MHA as an outcome variable, see Table 5). Sex, Health and Education were entered as predictors to account for alternative explanations for trends in the data. Age group, year of data collection and the interaction term (age group × year of data collection) were also entered as predictors: The age group label and the year of data collection were centralised by subtracting them form their own means, and then multiplied together to form the interaction predictor (Warner, 2008). For the DSST measure, all predictors were significant apart from health. A significant age group × year of data collection interaction term indicated that older adults' improvement in cognitive speed over time was more than young adults' improvement in cognitive speed over time, when also accounting for Sex, Health and Education and main effects. For the MHA measure all predictors were significant apart from the year of data collection predictor and the age group × year of data collection interaction predictor. Overall, these regressions indicate that age deficits in cognitive speed are diminishing over time, but that age differences in vocabulary are stable over time. However, as not all of the historic data included health and education measures some of the data were not included in the regressions. Using just three predictors – age group, year of data collection and age group × year of data collection, the crucial age group × year of data collection interaction term was significant for both speed as an outcome variable ($N = 1089$, $p = .008$) and for vocabulary as an outcome variable ($N = 1089$, $p = .003$). This indicates that older adults' cognition was improving over time for both speed and vocabulary, relative to young adults.

Discussion

Using historic data across seven years, this study showed that age deficits were diminishing over time and that the pattern was largely driven by improvements in older adults' cognition across time whilst young adults' cognition showed only minimal improvement across time. This matches results reported by Schaie (2012) who concluded that there are diminishing age deficits across time but still substantial age-related deficits for most abilities in the Seattle Longitudinal Study. The current result showing minimal improvement in young adults over time is also consistent with a reduction of the Flynn effect in modern society (Dutton, van der Linden and Lynn, 2016; Woodley of Menie et al., 2017; Wongupparaj et al., 2023). The diminishing age deficits observed in the current study were present for the measure of cognitive speed even when education, sex and health were accounted for in the regression model. This suggests that these basic control variables do not capture the full range of underlying factors responsible for time-based changes in age-related cognitive deficits.

General discussion

Three Studies reviewed independent sources of evidence to explore changes in cognition over time in older adults and the influence of time-based changes on age deficits in cognition. The overall message was that later-born older adults show higher cognitive performance than earlier-born older adults (Study 1 and Study 3) and that age deficits in cognition are diminishing over time (Study 2 and Study 3). Study 3 showed that changes in age deficits are largely driven by cognitive improvements in older adults over time, whilst cognitive performance in young adults over time shows minimal improvement.

Studies of the Flynn effect indicate a variety of underlying factors promoting the increase of cognitive performance over time such as improvements in education, nutrition, and health (see Wongupparaj et al., 2023, for a review), and the current data demonstrate an observation of the Flynn effect extending to include older adults' cognitive performance, similar to some existing accounts (e.g., Rönnlund & Nilsson, 2008, 2009). It has been hypothesized in studies showing a reduction of the Flynn effect that there is a limit to the cognitive benefits that can be derived from environmental factors (Dutton, van der Linden & Lynn, 2016; Woodley of Menie et al., 2017). The current results are consistent with these trends, it is argued here that as the Flynn effect approaches a limit in young adult populations, environmental advantages to young adults compared to older adults are diminishing along with a proportion of age-related cognitive deficits that might previously have been attributed to age-related decline. This is consistent with observations showing that a large proportion of cognitive ageing deficits are thought to be derived from the Flynn effect favouring the latter-born young adults in cross-sectional comparisons (Agbayani & Hiscock, 2013; Dickinson & Hiscock, 2010). The current argument is also consistent with reduced cognitive decline observed in longitudinal compared to cross-sectional data (e.g., Salthouse, 2009; Schaie, 2012). These results have both theoretical and applied implications for researchers and wider society.

The current results coincide with several papers noting a reduced incidence rates of dementia over time (Gao et al., 2019; Wu et al., 2017; Farina, Zhang et al., 2022). It is likely that undiagnosed dementia contributes partly to the assessment of age deficits in cognition as studies show that only around 40% of people with dementia are diagnosed (Aldus et al., 2020; Savva & Arthur, 2015) and dementia is often preceded by a period of mild cognitive impairment (Petersen, 2016). It is therefore possible that reduced dementia incidence rates over time contributed to the current result. The inverse is also possible – the current trend for reduced cognitive decline may have led to reduced dementia incidence rates in the same period. The underlying causes behind time-based trends in age differences in cognition and dementia are likely similar. Reductions in stroke incidence and heart disease were proposed as factors behind the trend

for reduced incidence rates of dementia (Gao et al., 2019), as well as reduced smoking and better education (Gao et al., 2019; Farina, Zhang et al., 2022; Wu et al., 2017).

Wu et al. (2017) noted that estimates of dementia trends over time are impeded by shifts in dementia awareness and clinical diagnostic thresholds. Establishing effective dementia diagnostic criteria is becoming increasingly pertinent with the advent of pharmacological treatments for the disease (e.g., McShane et al., 2019; Marucci et al., 2021) and the current article adds another factor for consideration in future estimates of pathological cognitive functionality in older adults. For example, rising IQ observed in the Flynn effect has resulted in the periodic re-norming of IQ tests used to measure Intellectual disability (Kanaya, Scullin & Ceci, 2003); therefore, in the light of the current findings, clinical definitions of cognitive impairment in older adults may also need to be periodically revised to aid effective dementia diagnosis.

A limitation of the current study and of the reviewed longitudinal research in Study 1 is that a definitive explanation for time-based trends in cognition is difficult to unpack (see also Rönnlund et al., 2005 & Salthouse, 2009, for a discussion of underlying factors linked to cross-sectional vs. longitudinal measures of age deficits). In most instances, where reported, education explained some of the time-based variance in cognition in the individual articles reviewed in Study 1, but education was not a significant predictor in the summarized data for Study 1 (see Table 1). In Study 3, education predicted some of the variance in cognitive scores after accounting for age and time period and it was clear in both Studies 1 and 3 that education was higher for later recruited cohorts of older adults. One limitation is that education is measured at a basic level, typically self-reported number of years, which might not capture sufficient variation to fully ascertain the role of education on time-based trends (e.g., education from parents and home environment; Wongupparaj et al., 2023). Nonetheless, education differences are one of the leading explanations utilized to explain the Flynn effect (Flynn, 1984). A variety of other changing environmental factors have also been hypothesized to account for the Flynn effect including nutrition, sanitation, vaccination, more cognitively stimulating environments, reduced exposure to pathogens (Wongupparaj et al., 2023) as well as economic conditions (e.g., Doblhammer,van den Berg & Fritze, 2013). These multiple underlying causes are consistent with the observation of a continuous increase in cognitive ability in Study 1 (see Fig. 1), with greater cognitive benefits shown between cohorts when those cohorts were separated further apart in time. Furthermore, in Study 2 it was shown that older adults' life expectancy was a good predictor of cognitive deficits, but this was difficult to dissociate from year of publication. Many of these explanatory factors are intercorrelated and future research should aim to identify the stronger positive environmental influences on population health and wellbeing to help clinicians and policymakers in allocation of resources.

A key limitation across the summarized research is that the samples are dominated by westernized populations, which is a common issue in psychology (Henrich, Heine, & Norenzayan; 2010) and in particular is an issue in cognitive psychology (Roberts et al., 2020). All of the large underlying longitudinal datasets contributing to Study 1 were from Europe or North America. None of the meta-analyses in Study 2 addressed or reported race or ethnicity information and neither did the underlying articles contributing to Study 3. In the current context this may be particularly pertinent given that the rise in IQ over time is hypothesized to be related to development in education and healthcare availability, and the Flynn effect is more evident in countries with initially lower performance levels (Hessel,et al., 2018; Wongupparaj et al., 2015). Moving forwards, it will be useful for researchers in this field to report racial demographics of their samples and take other steps towards racial inclusivity throughout the publication process as outlined by Roberts et al. (2020). Ongoing research in this field should also focus on a wider array of longitudinal data from non-western nations.

What factors may be unique to the young adult samples of cross-sectional comparisons? Definitive explanations for a reduced/negative Flynn effect in young adults have also been inconclusive but factors such as maternal age, immigration and dysgenics have been explored (Dutton, van der Linden & Lynn, 2016). Another trend in society is the rise of depression, which could also explain the current results in Studies 2 and 3. Whilst noted earlier that depression has been rising in older adults over time, the rise in young adults over time has been steeper (Weinberger et al., 2018), and depression is known to impede cognitive performance (McDermott & Ebmeier, 2009; Oi, 2017). Finally, young adult samples in psychological research are predominantly recruited from undergraduate cohorts, and with growth in education, such samples will naturally become representative of the average cognitive ability of the population, as opposed to representing individuals with higher ability.

In summary, the current data indicates how time-based cognitive changes may influence academics and society. It is now evidenced that cognitive deficits are likely becoming smaller and that academics in cognitive ageing should revise their expectations and sample sizes when studying age deficits in cognition. At a personal and public interest level, the current data also build upon evidence that cognitive decline may be smaller than previously thought as existing assessment of age-related cognitive deficits were likely influenced by environmental differences such as changes in education. Finally, the current data also show that older adults' cognition is improving over time which is a positive message for individuals but has implications for establishing pathological ageing when diagnosing dementia.

Supplementary information

Supplementary information is available for this paper here (https://doi.org/10.17605/OSF.IO/UAK7P).

Declaration of competing interest

The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data corresponding to this article can be found here: https://doi.org/10.17605/OSF.IO/UAK7P

Acknowledgements

The author thanks Dr Gary Jones for comments on an earlier version of this manuscript.

Appendix A

See Table A1.

References

Data from marked articles contribute to Study 1 (*), Study 2 (**), and Study 3(***).

*Ahrenfeldt, L. J., Lindahl-Jacobsen, R., Rizzi, S., Thinggaard, M., Christensen, K., & Vaupel, J. W. (2018). Comparison of cognitive and physical functioning of Europeans in 2004–05 and 2013. *International Journal of Epidemiology, 47*(5), 1518–1528. https://doi.org/10.1093/ije/dyy094

Agbayani, K. A., & Hiscock, M. (2013). Age-related change in Wechsler IQ norms after adjustment for the Flynn effect: Estimates from three computational models. *Journal of Clinical and Experimental Neuropsychology, 35*(6), 642–654. https://doi.org/10.1080/13803395.2013.806650

Aldus, C., Arthur, A., Dennington-Price, A., Millac, P., Richmond, P., Fox, C., … Savva, G. (2020). Undiagnosed dementia in primary care: A record linkage study. *Health Services and Delivery Research, 8*(20). https://doi.org/10.3310/hsdr08200

*Alwin, D. F., & McCammon, R. J. (2001). Aging, cohorts, and verbal ability. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 56*(3), S151–S161. https://doi.org/10.1093/geronb/56.3.S151

Armstrong, B. A., Ein, N., Wong, B. I., Gallant, S. N., & Li, L. (2021). Retracted: The effect of bilingualism on older adults' inhibitory control: a meta-analysis. *The Gerontologist, 61*(3), e102–e117. https://doi.org/10.1093/geront/gnz086

Badham, S. P.. The Older Population is More Cognitively able than in the Past and Age-related Deficits in Cognition are Diminishing Over Time. Data Set. https://osf.io/uak7p/.

***Badham, S. P., & Maylor, E. A. (2011). Age-related associative deficits are absent with nonwords. *Psychology and Aging, 26*, 689–694.

***Badham, S. P., & Maylor, E. A. (2012). Age-related associative deficits and the isolation effect. *Aging, Neuropsychology, and Cognition, 20*, 405–428.

***Badham, S. P., Estes, Z., & Maylor, E. A. (2012). Integrative and semantic relations equally alleviate age-related associative memory deficits. *Psychology and Aging, 27*, 141–152.

***Badham, S. P., Wade, K. A., Watts, H. J. E., Woods, N. G., & Maylor, E. A. (2013). Replicating distinctive facial features in lineups: identification performance in young versus older adults. *Psychonomic Bulletin & Review, 20*, 289–295.

***Badham, S. P., & Maylor, E. A. (2016). Antimnemonic effects of schemas in young and older adults. *Aging, Neuropsychology, and Cognition, 23*, 78–102.

***Badham, S. P., Poirier, M., Gandhi, N., Hadjivassiliou, A., & Maylor, E. A. (2016). Aging and memory as discrimination: Influences of encoding specificity, cue overload, and prior knowledge. *Psychology and Aging, 31*, 758–770.

***Badham, S. P., Sanborn, A. N., & Maylor, E. A. (2017). Deficits in category learning in older adults: rule-based versus clustering accounts. *Psychology and Aging, 32*, 473–488.

***Badham, S. P., Whitney, C., Sanghera, S., & Maylor, E. A. (2017). Word frequency influences on the list length effect and associative memory in young and older adults. *Memory, 25*, 816–830.

**Badham, S. P., Justice, L. V., Jones, L. N., & Myers, J. A. (2022). An older adult advantage in autobiographical recall. *Aging, Neuropsychology, and Cognition*, 1–27.

*Bancks, M., Alonso, A., Allen, N., Yaffe, K., & Carnethon, M. (2019). Temporal trends in cognitive function of older US adults associated with population changes in demographic and cardiovascular profiles. *Journal of Epidemiology and Community Health, 73*(7), 612–618. https://doi.org/10.1136/jech-2018-211985

*Beller, J., Kuhlmann, B. G., Sperlich, S., & Geyer, S. (2022). Secular improvements in cognitive aging: Contribution of education, health, and routine activities. *Journal of Aging and Health, 34*(6–8), 807–817. https://doi.org/10.1177/08982643211065571

*Ben-David, B. M., Erel, H., Goy, H., & Schneider, B. A. (2015). "Older is always better": Age-related differences in vocabulary scores across 16 years. *Psychology and Aging, 30*(4), 856–862. https://doi.org/10.1037/pag0000051

*Bloomberg, M., Dugravot, A., Dumurgier, J., Kivimaki, M., Fayosse, A., Steptoe, A., … Sabia, S. (2021). Sex differences and the role of education in cognitive ageing: Analysis of two UK-based prospective cohort studies. *The Lancet Public Health, 6*(2), e106–e115. https://doi.org/10.1016/S2468-2667(20)30258-9

Bopp, K. L., & Verhaeghen, P. (2020). Aging and n-back performance: A meta-analysis. *The Journals of Gerontology: Series B, Psychological Sciences and Social Sciences, 75*(2), 229–240. https://doi.org/10.1093/geronb/gby024

*Bordone, V., Scherbov, S., & Steiber, N. (2015). Smarter every day: The deceleration of population ageing in terms of cognition. *Intelligence, 52*, 90–96. https://doi.org/10.1016/j.intell.2015.07.005

*Brailean, A., Huisman, M., Prince, M., Prina, A. M., Deeg, D. J., & Comijs, H. (2018). Cohort differences in cognitive aging in the longitudinal aging study Amsterdam. *The Journals of Gerontology: Series B, Psychological Sciences and Social Sciences, 73*(7), 1214–1223. https://doi.org/10.1093/geronb/gbw129

Brunton, R. J., & Scott, G. (2015). Do we fear ageing? A multidimensional approach to ageing anxiety. *Educational Gerontology, 41*(11), 786–799. https://doi.org/10.1080/03601277.2015.1050870

Cahn-Weiner, D. A., Boyle, P. A., & Malloy, P. F. (2002). Tests of executive function predict instrumental activities of daily living in community-dwelling older individuals. *Applied Neuropsychology, 9*(3), 187–191. https://doi.org/10.1207/S15324826AN0903_8

**Cervera-Crespo, T., & González-Alvarez, J. (2017). Age and semantic inhibition measured by the Hayling task: A meta-analysis. *Archives of Clinical Neuropsychology, 32*(2), 198–214. https://doi.org/10.1093/arclin/acw088

Charles, S. T., & Carstensen, L. L. (2010). Social and emotional aging. *Annual Review of Psychology, 61*, 383–409. https://doi.org/10.1146/annurev.psych.093008.100448

Chen, X. J., Wang, Y., Liu, L. L., Cui, J. F., Gan, M. Y., Shum, D. H., & Chan, R. C. (2015). The effect of implementation intention on prospective memory: A systematic and meta-analytic review. *Psychiatry Research, 226*(1), 14–22. https://doi.org/10.1016/j.psychres.2015.01.011

Chen, F. T., Etnier, J. L., Chan, K. H., Chiu, P. K., Hung, T. M., & Chang, Y. K. (2020). Effects of exercise training interventions on executive function in older adults: A systematic review and meta-analysis. *Sports Medicine, 50*(8), 1451–1467. https://doi.org/10.1007/s40279-020-01292-x

*Choi, H., Schoeni, R. F., Martin, L. G., & Langa, K. M. (2018). Trends in the prevalence and disparity in cognitive limitations of Americans 55–69 years old. *The Journals of Gerontology: Series B, 73*(suppl_1), S29–S37. https://doi.org/10.1093/geronb/gbx155

*Christensen, K., Thinggaard, M., Oksuzyan, A., Steenstrup, T., Andersen-Ranberg, K., Jeune, B., … Vaupel, J. W. (2013). Physical and cognitive functioning of people older than 90 years: A comparison of two Danish cohorts born 10 years apart. *The Lancet, 382*(9903), 1507–1513. https://doi.org/10.1016/S0140-6736(13)60777-1

Dasgupta, P., Henshaw, C., Youlden, D. R., Clark, P. J., Aitken, J. F., & Baade, P. D. (2020). Global trends in incidence rates of primary adult liver cancers: A systematic review and meta-analysis. *Frontiers in Oncology, 10*, 171. https://doi.org/10.3389/fonc.2020.00171

Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., … Starr, J. M. (2009). Age-associated cognitive decline. *British Medical Bulletin, 92*(1), 135–152. https://doi.org/10.1093/bmb/ldp033

*Degen, C., Frankenberg, C., Toro, P., & Schröder, J. (2022). Differences in cognitive functioning in two birth cohorts born 20 years apart: data from the interdisciplinary longitudinal study of ageing. *Brain Sciences, 12*(2), 271. https://doi.org/10.3390/brainsci12020271

*de Rotrou, J., Wu, Y. H., Mabire, J. B., Moulin, F., de Jong, L. W., Rigaud, A. S., … Vidal, J. S. (2013). Does cognitive function increase over time in the healthy elderly? *PloS One, 8*(11), e78646.

Dickinson, M. D., & Hiscock, M. (2010). Age-related IQ decline is reduced markedly after adjustment for the Flynn effect. *Journal of Clinical and Experimental Neuropsychology, 32*(8), 865–870. https://doi.org/10.1080/13803391003596413

*Doblhammer, G., van den Berg, G. J., & Fritze, T. (2013). Economic conditions at the time of birth and cognitive abilities late in life: evidence from Ten European Countries. *PloS ONE, 8*(9), e74915.

*Dodge, H. H., Zhu, J., Lee, C. W., Chang, C. C. H., & Ganguli, M. (2014). Cohort effects in age-associated cognitive trajectories. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences, 69*(6), 687–694. https://doi.org/10.1093/Gerona/glt181

*Dodge, H. H., Zhu, J., Hughes, T. F., Snitz, B. E., Chang, C. C. H., Jacobsen, E. P., & Ganguli, M. (2017). Cohort effects in verbal memory function and practice effects: A population-based study. *International Psychogeriatrics, 29*(1), 137–148. https://doi.org/10.1017/S1041610216001551

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

Erickson, K. I., & Kramer, A. F. (2009). Aerobic exercise effects on cognitive and neural plasticity in older adults. *British Journal of Sports Medicine, 43*(1), 22–24. https://doi.org/10.1136/bjsm.2008.052498

**Erickson, W. B., Lampinen, J. M., & Moore, K. N. (2016). Eyewitness identifications by older and younger adults: A meta-analysis and discussion. *Journal of Police and Criminal Psychology, 31*(2), 108–121. https://doi.org/10.1007/s11896-015-9176-3

Farina, F. R., Bennett, M., Griffith, J. W., & Lenaert, B. (2022). Fear of memory loss predicts increased memory failures and lower quality of life in older adults: Preliminary findings from a fear-avoidance of memory loss (FAM) scale. *Aging & Mental Health, 26*(3), 486–492. https://doi.org/10.1080/13607863.2020.1856780

Farina, M. P., Zhang, Y. S., Kim, J. K., Hayward, M. D., & Crimmins, E. M. (2022). Trends in dementia prevalence, incidence, and mortality in the United States (2000–2016). *Journal of Aging and Health, 34*(1), 100–108. https://doi.org/10.1177/08982643211029716

Feart, C., Samieri, C., Rondeau, V., Amieva, H., Portet, F., Dartigues, J. F., … Barberger-Gateau, P. (2009). Adherence to a Mediterranean diet, cognitive decline, and risk of dementia. *JAMA, 302*(6), 638–648. https://doi.org/10.1001/jama.2009.1146

*Finkel, D., Reynolds, C. A., McArdle, J. J., & Pedersen, N. L. (2007). Cohort differences in trajectories of cognitive aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 62*(5), P286–P294. https://doi.org/10.1037/0033-2909.95.1.29

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

*Fox, M. C., & Mitchum, A. L. (2014). Confirming the cognition of rising scores: Fox and Mitchum (2013) predicts violations of measurement invariance in series completion between age-matched cohorts. *Plos One, 9*(5), e95780. https://doi.org/10.1371/journal.pone.0095780

**Fraundorf, S. H., Hourihan, K. L., Peters, R. A., & Benjamin, A. S. (2019). Aging and recognition memory: A meta-analysis. *Psychological Bulletin, 145*(4), 339. https://doi.org/10.1037/bul0000185

*Frith, E., & Loprinzi, P. D. (2019). 15-year secular trends in cognitive function among older adults in the United States. *Psychological Reports, 122*(3), 841–852. https://doi.org/10.1177/0033294118765227

*Fritze, T., Doblhammer, G., & van den Berg, G. J. (2014). Can individual conditions during childhood mediate or moderate the long-term cognitive effects of poor economic environments at birth? *Social Science & Medicine, 119*, 240–248. https://doi.org/10.1016/j.socscimed.2014.07.011

Gao, S., Burney, H. N., Callahan, C. M., Purnell, C. E., & Hendrie, H. C. (2019). Incidence of dementia and Alzheimer disease over time: A meta-analysis. *Journal of the American Geriatrics Society, 67*(7), 1361–1369. https://doi.org/10.1111/jgs.16027

*Gerstorf, D., Hülür, G., Drewelies, J., Eibich, P., Duezel, S., Demuth, I., … Lindenberger, U. (2015). Secular changes in late-life cognition and well-being: Towards a long bright future with a short brisk ending? *Psychology and Aging, 30*(2), 301. https://doi.org/10.1037/pag0000016

*Gerstorf, D., Ram, N., Hoppmann, C., Willis, S. L., & Schaie, K. W. (2011). Cohort differences in cognitive aging and terminal decline in the Seattle Longitudinal Study. *Developmental Psychology, 47*(4), 1026. https://doi.org/10.1037/a0023426

*Gerstorf, D., Ram, N., Drewelies, J., Duezel, S., Eibich, P., Steinhagen-Thiessen, E., … Ghisletta, P. (2023). Today's older adults are cognitively fitter than older adults were 20 years ago, but when and how they decline is no different than in the past. *Psychological Science, 34*(1), 22–34. https://doi.org/10.1177/09567976221118541

*Geyer, S., Kuhlmann, B. G., Beller, J., & Grasshoff, J. (2023). The role of school education in time-dependent changes of cognitive abilities in cohorts from midlife to old age. *Aging & Mental Health, 27*(4), 729–735. https://doi.org/10.1080/13607863.2022.2068132

*Graves, L. V., Drozdick, L., Courville, T., Farrer, T. J., Gilbert, P. E., & Delis, D. C. (2021). Cohort differences on the CVLT-II and CVLT3: Evidence of a negative Flynn effect on the attention/working memory and learning trials. *The Clinical Neuropsychologist, 35*(3), 615–632. https://doi.org/10.1080/13854046.2019.1699605

*Grasset, L., Jacqmin-Gadda, H., Proust-Lima, C., Pérès, K., Amieva, H., Dartigues, J. F., & Helmer, C. (2018). Temporal trends in the level and decline of cognition and disability in an elderly population: The PAQUID study. *American Journal of Epidemiology, 187*(10), 2168–2176. https://doi.org/10.1093/aje/kwy118

*Grasshoff, J., Beller, J., Kuhlmann, B. G., & Geyer, S. (2021). Increasingly capable at the ripe old age? Cognitive abilities from 2004 to 2013 in Germany, Spain, and Sweden. *Plos One, 16*(7), e0254038.

Groenewegen, A., Rutten, F. H., Mosterd, A., & Hoes, A. W. (2020). Epidemiology of heart failure. *European Journal of Heart Failure, 22*(8), 1342–1356. https://doi.org/10.1002/ejhf.1858

Gui, W. J., Li, H. J., Guo, Y. H., Peng, P., Lei, X., & Yu, J. (2017). Age-related differences in sleep-based memory consolidation: A meta-analysis. *Neuropsychologia, 97*, 46–55. https://doi.org/10.1016/j.neuropsychologia.2017.02.001

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation, 22*, 193–225. https://doi.org/10.1016/S0079-7421(08)60041-9

**Hayes, G. S., McLennan, S. N., Henry, J. D., Phillips, L. H., Terrett, G., Rendell, P. G., … Labuschagne, I. (2020). Task characteristics influence facial emotion recognition age-effects: A meta-analytic review. *Psychology and Aging, 35*(2), 295. https://doi.org/10.1037/pag0000441

*Henchoz, Y., Büla, C., von Gunten, A., Blanco, J. M., Seematter-Bagnoud, L., Démonet, J. F., … Santos-Eggimann, B. (2020). Trends in physical and cognitive performance among community-dwelling older adults in Switzerland. *The Journals of Gerontology: Series A, 75*(12), 2347–2353. https://doi.org/10.1093/Gerona/glaa008

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature, 466*(7302), 29.

*Hessel, P., Kinge, J. M., Skirbekk, V., & Staudinger, U. M. (2018). Trends and determinants of the Flynn effect in cognitive functioning among older individuals in 10 European countries. *Epidemiology and Community Health, 72*(5), 383–389. https://doi.org/10.1136/jech-2017-209979

*Hülür, G., Infurna, F. J., Ram, N., & Gerstorf, D. (2013). Cohorts based on decade of death: No evidence for secular trends favoring later cohorts in cognitive aging and terminal decline in the AHEAD study. *Psychology and Aging, 28*(1), 115–127. https://doi.org/10.1037/a0029965

*Hülür, G., Ram, N., Willis, S. L., Schaie, K. W., & Gerstorf, D. (2019). Cohort differences in cognitive aging: The role of perceived work environment. *Psychology and Aging, 34*(8), 1040–1054. https://doi.org/10.1037/pag0000355
Huo, L., Zhu, X., Zheng, Z., Ma, J., Ma, Z., Gui, W., & Li, J. (2021). Effects of transcranial direct current stimulation on episodic memory in older adults: A meta-analysis. *The Journals of Gerontology: Series B, 76*(4), 692–702. https://doi.org/10.1093/geronb/gbz130
Jaul, E., & Barron, J. (2017). Age-related diseases and clinical and public health implications for the 85 years old and over population. *Frontiers in Public Health, 5*, 335. https://doi.org/10.3389/fpubh.2017.00335
*Johnsen, B., Strand, B. H., Martinaityte, I., Mathiesen, E. B., & Schirmer, H. (2021). Improved cognitive function in the Tromso study in Norway from 2001 to 2016. *Neurology: Clinical Practice, 11*(6), e856–e866. https://doi.org/10.1212/CPJ.0000000000001115
Kanaya, T., Scullin, M. H., & Ceci, S. J. (2003). The Flynn effect and U.S. policies: The impact of rising IQ scores on American Society via mental retardation diagnoses. *American Psychologist, 58*(10), 778–790. https://doi.org/10.1037/0003-066X.58.10.778
*Karlsson, P., Thorvaldsson, V., Skoog, I., Gudmundsson, P., & Johansson, B. (2015). Birth cohort differences in fluid cognition in old age: Comparisons of trends in levels and change trajectories over 30 years in three population-based samples. *Psychology and Aging, 30*(1), 83–94. https://doi.org/10.1037/a0038643
Kingston, A., Wohland, P., Wittenberg, R., Robinson, L., Brayne, C., Matthews, F. E., … Weller, R. (2017). Is late-life dependency increasing or not? A comparison of the Cognitive Function and Ageing Studies (CFAS). *The Lancet, 390*(10103), 1676–1684. https://doi.org/10.1016/S0140-6736(17)31575-1
*Lewis, K. L., & Zelinski, E. M. (2010). List and text recall differ in their predictors: Replication over samples and time. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 65*(4), 449–458. https://doi.org/10.1093/geronb/gbq034
Liu, Z., Suo, C., Mao, X., Jiang, Y., Jin, L., Zhang, T., & Chen, X. (2020). Global incidence trends in primary liver cancer by age at diagnosis, sex, region, and etiology, 1990–2017. *Cancer, 126*(10), 2267–2278. https://doi.org/10.1002/cncr.32789
*Llewellyn, D. J., Matthews, F. E., & The Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). (2009). Increasing levels of semantic verbal fluency in elderly English adults. *Aging, Neuropsychology, and Cognition, 16*(4), 433–445. https://doi.org/10.1080/13825580902773867
Lövdén, M., Fratiglioni, L., Glymour, M. M., Lindenberger, U., & Tucker-Drob, E. M. (2020). Education and cognitive functioning across the life span. *Psychological Science in the Public Interest, 21*(1), 6–41.
*Luo, L., & Warren, J. R. (2023). Describing and explaining age, period, and cohort trends in Americans' vocabulary knowledge. *Population Research and Policy Review, 42*(3), 31. https://doi.org/10.1007/s11113-023-09771-5
Mathers, C. D., Stevens, G. A., Boerma, T., White, R. A., & Tobias, M. I. (2015). Causes of international increases in older age life expectancy. *The Lancet, 385*(9967), 540–548. https://doi.org/10.1016/S0140-6736(14)60569-9
Madsen, T. E., Khoury, J. C., Leppert, M., Alwell, K., Moomaw, C. J., Sucharew, H., … Kleindorfer, D. O. (2020). Temporal trends in stroke incidence over time by sex and age in the CNKSS. *Stroke, 51*(4), 1070–1076. https://doi.org/10.1161/STROKEAHA.120.028910
Martyr, A., & Clare, L. (2012). Executive function and activities of daily living in Alzheimer's disease: A correlational meta-analysis. *Dementia and Geriatric Cognitive Disorders, 33*(2–3), 189–203. https://doi.org/10.1159/000338233
Marucci, G., Buccioni, M., Dal Ben, D., Lambertucci, C., Volpini, R., & Amenta, F. (2021). Efficacy of acetylcholinesterase inhibitors in Alzheimer's disease. *Neuropharmacology, 190*, Article 108352. https://doi.org/10.1016/j.neuropharm.2020.108352
McShane, R., Westby, M. J., Roberts, E., Minakaran, N., Schneider, L., Farrimond, L. E., Maayan, N., Ware, J., & Debarros, J. (2019). Memantine for dementia. *Cochrane Database of Systematic Reviews*, (3), Article CD003154. https://doi.org/10.1002/14651858.CD003154.pub6
McDermott, L. M., & Ebmeier, K. P. (2009). A meta-analysis of depression severity and cognitive function. *Journal of Affective Disorders, 119*(1–3), 1–8. https://doi.org/10.1016/j.jad.2009.04.022
*Munukka, M., Koivunen, K., von Bonsdorff, M., Sipilä, S., Portegijs, E., Ruoppila, I., & Rantanen, T. (2021). Birth cohort differences in cognitive performance in 75- and 80-year-olds: a comparison of two cohorts over 28 years. *Aging Clinical and Experimental Research, 33*, 57–65. https://doi.org/10.1007/s40520-020-01702-0
*Nielsen, C. R., Ahrenfeldt, L. J., Jeune, B., Christensen, K., & Lindahl-Jacobsen, R. (2022). Development in life expectancy with good and poor cognitive function in the elderly European Population from 2004–05 to 2015. *European Journal of Epidemiology, 37*(5), 495–502. https://doi.org/10.1007/s10654-022-00860-x
Nyberg, L., Magnussen, F., Lundquist, A., Baaré, W., Bartrés-Faz, D., Bertram, L., … Fjell, A. M. (2021). Educational attainment does not influence brain aging. *Proceedings of the National Academy of Sciences, 118*(18).
*Oi, K. (2017). Inter-connected trends in cognitive aging and depression: Evidence from the health and retirement study. *Intelligence, 63*, 56–65. https://doi.org/10.1016/j.intell.2017.05.004
Oberste, M., Sharma, S., Bloch, W., & Zimmer, P. (2021). Acute exercise-induced set shifting benefits in healthy adults and its moderators: A systematic review and meta-analysis. *Frontiers in Psychology, 12*, Article 528352. https://doi.org/10.3389/fpsyg.2021.528352
Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: a meta-analysis. *Psychology and Aging, 23*(1), 104–118. https://doi.org/10.1037/0882-7974.23.1.104
*Overton, M., Pihlsgård, M., & Elmståhl, S. (2018). Up to speed: Birth cohort effects observed for speed of processing in older adults: Data from the Good Ageing in Skåne population study. *Intelligence, 67*, 33–43. https://doi.org/10.1016/j.intell.2018.01.002
Park, D. C., Polk, T. A., Mikels, J. A., Taylor, S. F., & Marshuetz, C. (2022). Cerebral aging: Integration of brain and behavioral models of cognitive function. *Dialogues in Clinical Neuroscience, 3*(3), 151–165. https://doi.org/10.31887/DCNS.2001.3.3/dcpark
Raven, J. C., Raven, J., & Court, J. H. (1988). *The Mill Hill vocabulary scale*. London, England: H. K. Lewis.
Petersen, R. C. (2016). Mild cognitive impairment. CONTINUUM: Lifelong Learning. *Neurology, 22*(2 Dementia), 404–418. https://doi.org/10.1212/CON.0000000000000313
**Rey-Mermet, A., & Gade, M. (2018). Inhibition in aging: What is preserved? What declines? A meta-analysis. *Psychonomic Bulletin & Review, 25*(5), 1695–1716. https://doi.org/10.3758/s13423-017-1384-7
**Rhodes, S., Greene, N. R., & Naveh-Benjamin, M. (2019). Age-related differences in recall and recognition: A meta-analysis. *Psychonomic Bulletin & Review, 26*(5), 1529–1547. https://doi.org/10.3758/s13423-019-01649-y
Roberts, K. L., & Allen, H. A. (2016). Perception and cognition in the ageing brain: A brief review of the short-and long-term links between perceptual and cognitive decline. *Frontiers in aging neuroscience, 8*, 39.
Roberts, S. O., Bareket-Shavit, C., Dollins, F. A., Goldie, P. D., & Mortenson, E. (2020). Racial inequality in psychological research: Trends of the past and recommendations for the future. *Perspectives on Psychological Science, 15*(6), 1295–1309.
*Rodgers, W. L., Ofstedal, M. B., & Herzog, A. R. (2003). Trends in scores on tests of cognitive ability in the elderly US population, 1993–2000. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 58*(6), S338–S346. https://doi.org/10.1093/geronb/58.6.S338
Rönnlund, M., & Nilsson, L. G. (2008). The magnitude, generality, and determinants of Flynn effects on forms of declarative memory and visuospatial ability: Time-sequential analyses of data from a Swedish cohort study. *Intelligence, 36*(3), 192–209.
Rönnlund, M., & Nilsson, L. G. (2009). Flynn effects on sub-factors of episodic and semantic memory: Parallel gains over time and the same set of determining factors. *Neuropsychologia, 47*(11), 2174–2180.
Rönnlund, M., Nyberg, L., Bäckman, L., & Nilsson, L. G. (2005). Stability, growth, and decline in adult life span development of declarative memory: cross-sectional and longitudinal data from a population-based study. *Psychology and Aging, 20*(1), 3–18.
Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403
Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging, 30*(4), 507–514. https://doi.org/10.1016/j.neurobiolaging.2008.09.023
Salmon, D. P., & Bondi, M. W. (2009). Neuropsychological assessment of dementia. *Annual Review of Psychology, 60*, 257–282. https://doi.org/10.1146/annurev.psych.57.102904.190024
Savva, G. M., & Arthur, A. (2015). Who has undiagnosed dementia? A cross-sectional analysis of participants of the Aging, Demographics and Memory Study. *Age and Ageing, 44*(4), 642–647. https://doi.org/10.1093/ageing/afv020

Schafer, E. C., Miller, S., Manning, J., Zhang, Q., Lavi, A., Bodish, E., … De Blaey, E. (2021). Meta-analysis of speech recognition outcomes in younger and older adults with cochlear implants. *American Journal of Audiology, 30*(3), 481–496. https://doi.org/10.1044/2021_AJA-20-00141
*Schaie, K. W. (2012). Developmental Influences on Adult Intelligence: The Seattle Longitudinal Study; 2nd ed. Oxford University Press: Oxford. https://doi.org/10.1093/acprof:osobl/9780195386134.001.0001.
Seblova, D., Berggren, R., & Lövdén, M. (2020). Education and age-related decline in cognitive performance: Systematic review and meta-analysis of longitudinal cohort studies. *Ageing Research Reviews, 58*, Article 101005.
*Skirbekk, V., Stonawski, M., Bonsang, E., & Staudinger, U. M. (2013). The Flynn effect and population aging. *Intelligence, 41*(3), 169–177. https://doi.org/10.1016/j.intell.2013.02.001
*Steiber, N. (2015). Population aging at cross-roads: Diverging secular trends in average cognitive functioning and physical health in the older population of Germany. *PloS One, 10*(8), e0136583.
Stephan, Y., Sutin, A. R., Luchetti, M., Aschwanden, D., & Terracciano, A. (2021). Subjective age and verbal fluency among middle aged and older adults: A meta-analysis of five cohorts. *Archives of Gerontology and Geriatrics, 97*, Article 104527. https://doi.org/10.1016/j.archger.2021.104527
Stephan, B., Tang, E. Y., Pakpahan, E., Biswas, B., Gupta, A., Fairley, A., … Siervo, M. (2022). Secular trends in dementia free cognitive function in older adults: A systematic review. *Journal of Alzheimer's Disease, 88*(2), 417–428. https://doi.org/10.3233/JAD-220162
*Tampubolon, G. (2015). Cognitive ageing in Great Britain in the new century: Cohort differences in episodic memory. *PloS One, 10*(12), e0144907.
**Theisen, M., Lerche, V., von Krause, M., & Voss, A. (2021). Age differences in diffusion model parameters: A meta-analysis. *Psychological Research, 85*(5), 2012–2021. https://doi.org/10.1007/s00426-020-01371-8
*Thorvaldsson, V., Karlsson, P., Skoog, J., Skoog, I., & Johansson, B. (2017). Better cognition in new birth cohorts of 70 year olds, but greater decline thereafter. *The Journals of Gerontology: Series B, Psychological Sciences and Social Sciences, 72*(1), 16–24. https://doi.org/10.1093/geronb/gbw125
*Turcotte, V., Potvin, O., Dadar, M., Hudon, C., Duchesne, S., & Initiative, A. D. N. (2022). Birth cohorts and cognitive reserve influence cognitive performances in older adults. *Journal of Alzheimer's Disease, 85*(2), 587–604. https://doi.org/10.3233/JAD-215044
United Nations, Department of Economic and Social Affairs, Population Division (2022). World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/NO. 3.
**Vallesi, A., Tronelli, V., Lomi, F., & Pezzetta, R. (2021). Age differences in sustained attention tasks: a meta-analysis. *Psychonomic Bulletin & Review, 28*(6), 1755–1775. https://doi.org/10.3758/s13423-021-01908-x
Verhaeghen, P. (2003). Aging and vocabulary score: A meta-analysis. *Psychology and Aging, 18*(2), 332–339.
*Vonk, J. M., Rentería, M. A., Avila, J. F., Schupf, N., Noble, J. M., Mayeux, R., … Manly, J. J. (2019). Secular trends in cognitive trajectories of diverse older adults. *Alzheimer's & Dementia, 15*(12), 1576–1587. https://doi.org/10.1016/j.jalz.2019.06.4944
Wallhagen, M. I., Strawbridge, W. J., Cohen, R. D., & Kaplan, G. A. (1997). An increasing prevalence of hearing impairment and associated risk factors over three decades of the Alameda County Study. *American Journal of Public Health, 87*(3), 440–442. https://doi.org/10.2105/AJPH.87.3.440
Wang, P., Liu, H. H., Zhu, X. T., Meng, T., Li, H. J., & Zuo, X. N. (2016). Action video game training for healthy adults: A meta-analytic study. *Frontiers in Psychology, 7*, 907. https://doi.org/10.3389/fpsyg.2016.00907
Ware, A. T., Kirkovski, M., & Lum, J. A. (2020). Meta-analysis reveals a bilingual advantage that is dependent on task and age. *Frontiers in Psychology, 11*, 1458. https://doi.org/10.3389/fpsyg.2020.01458
Warner, R. M. (2008). *Applied statistics: From bivariate through multivariate techniques*. Sage.
Wechsler, D. (1981). *Manual for the Wechsler adult intelligence scale, revised*. New York: Psychological Corporation.
Weinberger, A. H., Gbedemah, M., Martinez, A. M., Nash, D., Galea, S., & Goodwin, R. D. (2018). Trends in depression prevalence in the USA from 2005 to 2015: widening disparities in vulnerable groups. *Psychological Medicine, 48*(8), 1308–1315. https://doi.org/10.1017/S0033291717002781
Woodley of Menie, M. A., Peñaherrera, M., Fernandes, H. B. F., & Figueredo, A. J. (2017). What causes the anti-Flynn Effect? A data synthesis and analysis of predictors. *Evolutionary Behavioral Sciences, 12*(4), 276–295. https://doi.org/10.1037/ebs0000106
Wongupparaj, P., Kumari, V., & Morris, R. G. (2015). A Cross-Temporal Meta-Analysis of Raven's Progressive Matrices: Age groups and developing versus developed countries. *Intelligence, 49*, 1–9. https://doi.org/10.1016/j.intell.2014.11.008
Wongupparaj, P., Wongupparaj, R., Morris, R. G., & Kumari, V. (2023). Seventy years, 1000 samples, and 300,000 SPM scores: A new meta-analysis of Flynn effect patterns. *Intelligence, 98*, Article 101750. https://doi.org/10.1016/j.intell.2023.101750
Wu, Y. T., Beiser, A. S., Breteler, M., Fratiglioni, L., Helmer, C., Hendrie, H. C., … Brayne, C. (2017). The changing prevalence and incidence of dementia over time-current evidence. *Nature Reviews Neurology, 13*(6), 327–339. https://doi.org/10.1038/nrneurol.2017.63
*Yang, Y., & Land, K. C. (2006). A mixed models approach to the age-period-cohort analysis of repeated cross-section surveys, with an application to data on trends in verbal test scores. *Sociological Methodology, 36*(1), 75–97. https://doi.org/10.1111/j.1467-9531.2006.00175.x
*Zelinski, E. M., & Kennison, R. F. (2007). Not your parents' test scores: cohort reduces psychometric aging effects. *Psychology and Aging, 22*(3), 546–557. https://doi.org/10.1037/0882-7974.22.3.546
*Zheng, H. (2021). A new look at cohort trend and underlying mechanisms in cognitive functioning. *The Journals of Gerontology: Series B, Psychological Sciences and Social Sciences, 76*(8), 1652–1663. https://doi.org/10.1093/geronb/gbaa107
*Zhou, Y. (2022). Development over time in cognitive function among European 55–69-year-olds from 2006 to 2015, and differences of region, gender, and education. *Comparative Population Studies, 47*. https://doi.org/10.12765/CPoS-2022-04

Further reading

***Badham, S. P., Hay, M., Foxon, N., Kaur, K., & Maylor, E. A. (2016). When does prior knowledge disproportionately benefit older adults' memory? *Aging, Neuropsychology, and Cognition, 23*, 338–365.

Life's Rich Pageant: John Updike's 'Rabbit' Tetralogy

Ken Shea

'To desire to give verse-rhythm to prose, yet to leave it prose and very much prose and to write about ordinary life as histories and epics are written, yet without falsifying the subject, is perhaps an absurd idea. Sometimes I almost think it is. But it may also be a great experiment and very original.'

-Gustave Flaubert (excerpted from letter to Louise Colet, dated March 27th, 1853)

'A writer's standard of fidelity to the truth should be so high that his invention, out of his experience, should produce a truer account than anything factual can be.' -Ernest Hemingway

In an alliance possibly symbolic of the problems soon to be discussed, Stanford University and *Esquire* magazine collaborated in 1960 to host a topical literary symposium, ponderously billed as 'Writing in America Today.' One speaker at the event was an inimitable writer hailing from Newark, New Jersey: Philip Roth, who put a fictionalized Newark on the map to such an extent that there now exists, life imitating art, a Philip Roth Bus Tour of Newark. The year previous to the Stanford University speech, Philip Roth won the National Book Award for *Goodbye, Columbus*, which manages to 'ride the tiger' and roundly characterize social and political trends without drowning. The main predicament Roth perceived in the speech delivered at Stanford, 'Writing American Fiction', was the following: 'the American writer in the middle of the twentieth century has his hands full in trying to understand, describe, and then make *credible* much of American reality. It stupefies, it sickens, it infuriates, and finally it is even a kind of embarrassment to one's own meager imagination.' Roth, then, seems to address writers as a class more directly: 'The actuality is continually outdoing our talents, and the culture tosses up figures almost daily that are the envy of any novelist.' The preposterousness of American culture in the twentieth century, in other words, paradoxically rendered the writerly attempts at fiction obsolete, incomprehensible, beside the point, or too continuous with the falsehoods of commerce and politics to be properly appreciated.

> 'For a writer of fiction to feel that he does not really live in his own country - as represented by *Life* or by what he experiences when he steps out the front door - must seem a serious occupational impediment. For what will his subject be? His landscape? One would think that we might get a high proportion of historical novels or contemporary satire - or perhaps just nothing. No books. Yet almost weekly one finds on the best-seller list another novel which is set in Mamaroneck or New York City or Washington, with characters moving through a world of dishwashers and TV sets and advertising agencies and senatorial investigations. It all *looks* as though writers are turning out books about the world. There is *Cash McCall* and *The Man in the Gray Flannel Suit* and *Marjorie Morningstar* and *The Enemy Camp* and *Advise and Consent*, and so on. But what is noteworthy is that these books aren't very good.' -Philip Roth

Turning away from 'lesser minds and talents,' Roth mentions Norman Mailer, who blurred the distinction between writer and subject to the vanishing point with *Advertisements for Myself*. 'An infuriating, self-indulgent, boisterous, mean book, not much worse than most advertising we have to put up with - but, taken as a whole, curiously moving in its revelation of despair so great that the man who bears it, or is borne by it, seems for the time being to have given up on making an imaginative assault upon the American experience, and has instead

become the champion of a kind of public revenge,' according to Roth. The problem is that Mailer's project is more of an unrepeatable stunt than a way to consistently treat social problems in fiction. Even the metafiction Henry Miller made a stock in trade can become a victim of its own success when author and character fuse to create an epistemic closure around the novels themselves, arresting the notion of uncomplicated social commentary. In fact, the transgressive SF novelist J.G. Ballard [1] apprehended Henry Miller as 'A Working-class Proust', while conceding: 'Another problem Miller poses for the reader is that of knowing who exactly he is. Henry Miller is the hero of his fiction and his own greatest creation, so much so that, as [Henry Miller biographer] Robert Ferguson remarks, he became a hybrid of man and book. This conflation of author and text is anathema at present, given our desperate need to demystify the writer and shoulder aside any claims he may make to the ownership of his own text. Irritatingly for the deconstructionists, Miller's novels, like his life, only make sense in terms of the mythic illusions he managed to weave around them.' Philip Roth's frustration centers around 'the spurning of our [shared] world' in a fictional context, which 'another of our most gifted writers, Bernard Malamud,' succumbs to. Roth characterizes the 'baseball' occurring in Bernard Malamud's *The Natural* as a 'wild, wacky game', openly defying Newtonian physics, bearing only a passing resemblance to actual 'baseball as it is played in Yankee Stadium.' Roth [2] more or less closes the Stanford University speech by scratching his head at the unrepresentative, exuberant *deus ex machina* of a 'big, sloppy millionaire' named Eugene Henderson galloping around a field right after touching down in Newfoundland, a victory lap of sorts upon returning from an African adventure à la Saul Bellow's *Henderson the Rain King*. Rhetorically: 'And so we leave Henderson, a very happy man. Where? In the Arctic. The picture has stayed with me since I read the book a year ago: of a man who finds energy and joy in an imaginary Africa, and celebrates it on an unpeopled, ice-bound vastness.' The endings of J.D. Salinger's *The Catcher in the Rye* and Ralph Ellison's *The Invisible Man* were hardly less chilling in Roth's estimation.

---

[1] The novels compromising the 'urban trilogy' (*Crash*, *Concrete Island*, and *High-Rise*, published a year apart in 1973, 1974, and 1975, respectively) subvert bourgeois familiarity by exteriorizing drives/psychopathology. The main character in *Crash* is James Ballard. According to a self-assessment, the two chief influences on Ballard the author were the Surrealists (e.g., Max Ernst, Yves Tanguy, René Magritte, and Salvador Dali) and psychoanalysis. In a review titled 'The Coming of the Unconscious' ostensibly addressing *The History of Surrealist Painting* (Marcel Jean) and *Surrealism* (Patrick Walberg), Ballard finds: 'The images of surrealism are the iconography of inner space. Popularly regarded as a lurid manifestation of fantastic art concerned with states of dream and hallucination, surrealism is the first movement, in the words of Odilon Redon, to place 'the logic of the visible at the service of the invisible.' This calculated submission of the impulses and fantasies of our inner lives to the rigours of time and space, to the formal inquisition of the sciences, psychoanalysis preeminent among them, produces a heightened or alternate reality beyond that familiar to our sight or senses. What uniquely characterizes the fusion of the outer world of reality and the inner world of the psyche (which I have termed 'inner space') is its redemptive and therapeutic power. To move through these landscapes is a journey of return to one's innermost being' ('The Coming of the Unconscious', *A User's Guide to the Millenium*, page 84).

[2] In fairness, Philip Roth would go on to develop one of the most deep and riveting relationships to a literary alter ego, Nathan Zuckerman, perhaps ever conceived/rendered in the history of contemporary fiction writing. The 'fictional' writer Nathan Zuckerman, celebrated *ad nauseam* for writing (who could forget?) the salacious romp *Carnovsky*, self-evidently a stand-in for the unwanted (and unwonted) notoriety Roth received in the wake of *Portnoy's Complaint*, proactively wants to distance himself from the postmodern iconoclasm of writers like Thomas Pynchon and John Barth (fellow writers mentioned by name near the end of *Zuckerman Unbound*), mirroring Roth's own desire to craft a more representative, serious, and *credible* twentieth-century novel. In any case, the overwrought conflation in the public consciousness between imagination and autobiography regressed such that the flesh-and-blood Philip Roth was impelled to clarify: 'A novel in the guise of a confession [*Portnoy's Complaint*] was received and judged by any number of readers as a confession in the guise of a novel.' What's more, Philip Roth at one point - *talk about* straining credulity - was saddled with an unsolicited real-life Israeli impersonator, a bespectacled doppelganger possessed of a dubious political agenda, answering to 'Philip Roth,' which partly aroused the machinations in *Operation Shylock*: 'In this book (which may or may not be fiction), Philip Roth meets a man who may or may not be Philip Roth. Because someone with that name has been touring Israel, promoting a bizarre reverse exodus of the Jews. Roth is intent on stopping him, even if that means impersonating his own impersonator.' The actual Philip Roth was frequently accused of being a self-hating Jew - a charge Roth found patently ridiculous. Philip Roth, like Saul Bellow, has given a lot of thought to the interrelations between Jewishness, contemporary society, writing, and living in the United States. See the following essays by Roth: 'New Jewish Stereotypes', 'Writing About Jews', and 'Imagining Jews', collected in the Library of America's #300, a feast of Roth's non-fiction, *Philip Roth: Why Write?*

Cf. https://www.youtube.com/watch?v=5G8Z2vrLpzA

John Updike's 'Rabbit' tetralogy struck a resounding chord with the public and critics (the third and fourth installments, *Rabbit is Rich* and *Rabbit at Rest* each won a Pulitzer Prize) partly because the protagonist, Harry 'Rabbit' Angstrom, was on the same journey, more or less consciously, as the vast majority of twentieth-century readers. Rabbit's is a 'quest for a successful hedonism that would enable man to enjoy an otherwise meaningless life,' according to the author, John Updike. The novels comprising the tetralogy - *Rabbit, Run*, *Rabbit Redux*, *Rabbit is Rich*, and *Rabbit at Rest* - were published in 1960, 1971, 1981, and 1990, respectively, or about a decade apart from each other, allowing Updike to track and dramatize changes undergone in the country in real time, albeit in the fictionalized yet fully-fleshed out city of Brewer, Pennsylvania. (The backdrop for the first novel is said to be Mt. Judge, which is a suburb of Brewer, only the fifth-largest city in the state of Pennsylvania. The suburb of Mt. Judge functions to a degree like the fictionalized town of Gopher Prairie - modeled after Sauk Centre, Minnesota - in Sinclair Lewis's satirical novel *Main Street*. Sinclair Lewis apparently upset many locals (yokels?), who thought they could detect their likeness in Lewis's accomplished novel. Elsewhere, Updike did actually fictionalize Shillington, a modest Pennsylvania Dutch town and suburb of Reading where the author was born in 1932, in short stories where a place called 'Olinger' serves as a setting.) A close friend of John Updike's said at the time of *Rabbit, Run*'s publication that the character of Harry 'Rabbit' Angstrom is 'a picture of John, if he had been a better basketball player and had married a hometown girl.' The description, however, feels off, given Rabbit's flightiness, habits, interests, and more modest intellectual abilities - unlike *The Centaur*, winner of National Book Award for Fiction in 1964, being about John Updike's real father, a high school math teacher, and *Of the Farm*, published in 1965, possessing a strong-willed maternal character supposedly based on Updike's actual mother. John's more hardscrabble upbringing is intimated in *The Poorhouse Fair*, a novel originally published in 1959 steeped in memories of a Shillington upbringing. John Updike was the class president, editor of the school newspaper, and graduated with an straight-A average prior to receiving a full scholarship to Harvard University, where he graduated *summa cum laude*. Updike apparently banged out his first short story on a typewriter at the age of eight and had a poem published at the age of nine. Aside from short stories and poems, Updike preserved an early interest in art and received a Knox Fellowship to the Ruskin School of Drawing and Fine Art. The fellowship was to take place at Oxford but the occasion may have been overshadowed in Updike's mind by the acceptance of a short story and a poem by *The New Yorker*. Indeed, Updike called the event 'the ecstatic breakthrough of my literary life.' Clearly not content to rest on his laurels, Updike adopted a workingman's attitude to writing and rigorously followed a daily schedule: 'I see myself as keeping hours like the dentist although not as long,' Updike once reflected. One Updike biography credited the writer, all told, with '23 novels, 18 short story collections, 12 collections of poetry, 4 children's books, and 12 collections of non-fiction.' Close personal friend, Adam Gopnik, fellow long-time contributor to *The New Yorker*, spotlighted *Hugging the Shore* and *Picked-up Pieces* as reviews and essays, among Updike's non-fiction, to reach for. Sad to say but if Washington Irving (or rather Geoffrey Crayon, Jonathan Oldstyle, Launcelot Langstaff, and perhaps also Diedrich Knickerbocker) is said to be the nation's *first* man of letters (William Makepeace Thackeray anointed Irving 'the first ambassador whom the New World of Letters sent to the Old), then Updike might well be the *last* of a dying breed. In a similar vein, Gore Vidal is occasionally said to have been the last *public* intellectual active in the United States. [3][4][5]

Addressing a crowd in Stockholm, William Faulkner, associated with the Southern Gothic movement, considered the 'human heart in conflict with itself' as the *sine qua non* for 'good writing' or writing worth doing at all, in a fondly-remembered Nobel Prize speech. The protagonist, if that's the right word, of Updike's 'Rabbit' tetralogy, Harry 'Rabbit' Angstrom (the parts of the surname 'Ang' and 'strom' are suggestive of *angst* and storm in German), absolutely portrays such a conflicted individual. [6] The reader catches up with Harry working a lackluster job at five-and-dime stores - demonstrating the MagiPeeler™ - after 'serving' in the Korean War. (Rabbit did not, in fact, leave the state of Texas vis-à-vis potential deployment but managed to work in trips to a whorehouse while frequenting the Lone Star State.) Living in a two-family frame house back in Mt. Judge with a gone-to-seed, pregnant and heavy-drinking wife, Janice, and a two-year-old son, Nelson, Harry is a 26-year-old has-been high school basketball star. Harry will later confide in the local Episcopalian minister Jack Eccles, with whom Harry eventually has regular golfing tee times loosely designed around a confused, modern-day notion of pastoral counseling: 'I once played a game [basketball] real well. I really did. And after you're first-rate at something, no matter what, it kind of takes the kick out of being second-rate. And that little thing Janice and I had going, boy, it was really second-rate.' Harry is less than impressed with the life of bourgeois comfort and workaday routine being laid out before him, perhaps taking note of his own beaten-down linotypist father, and decides to make a break for it. Impulsively hatching a plan to drive down to Florida - essentially commandeering the family vehicle, come what may, for the jaunt - Rabbit eventually stalls, psychologically, in West Virginia and determines to head back to Pennsylvania, of course not without losing his bearings yet again. Rabbit, the name conjuring a givenness to flight, certain physical characteristics, the gratification of carnal appetites, and vigilance against being boxed in, will eventually toggle between Janice's uninspired abode and shacking up with figures from a less-troubled past, including Rabbit's former basketball coach, Tothero, and an erstwhile underclassman named Ruth, a part-time lady of the night - Updike's supporting characters are vividly drawn, there can be no doubt - who seems less taken with Harry than Harry is with himself. [7]

---

[3] Adam Gopnik remembers John Updike, author and friend: https://www.youtube.com/watch?v=lJSjvhCEZ-g

[4] J.D. Salinger's *The Catcher in the Rye* and Ralph Ellison's *The Invisible Man* can be found alongside John Updike's fast-moving *The Coup* (1978) and Mary McCarthy's underrated comic masterpiece *The Groves of Academe* in Anthony Burgess's *Ninety-Nine Novels: The Best in English Since 1939 - A Personal Choice*.

https://en.wikipedia.org/wiki/Ninety-nine_Novels

[5] John Updike was a far more diverse writer than detractors are willing (or able) to acknowledge. *The Coup* is a black comedy set in Sub-Saharan Africa, *The Centaur* is shot through with mythology, *Brazil* and *The Witches of Eastwick* accommodate plenty of magical realism, and *Gertrude and Claudius* owe a debt to William Shakespeare's *Hamlet*. The 1986 novel *Roger's Version* is another black comedy, about technology, and *S.* transmutes Hawthorne's *The Scarlet Letter*.

[6] Although the conventional wisdom is to term these books a 'tetralogy' (and not a pentalogy) - Everyman's Library published an omnibus of the four books in 1995, issue #214 - Updike wrote a related novella titled *Rabbit Remembered*, continuing the happy alliterative tendency, if not the life, *per se*, of the titular character.

[7] The name 'Angstrom' could also partly derive from physics: 'unit of measurement. angstrom (Å), unit of length, equal to 10−10 metre, or 0.1 nanometre. It is used chiefly in measuring wavelengths of light.'

‘Boys are playing basketball around a telephone pole with a backboard bolted to it. Legs, shouts. The scrape and snap of Keds on loose alley pebbles seems to catapult their voices high into the moist March air blue above the wires. Rabbit Angstrom, coming up the alley in a business suit, stops and watches, though he’s twenty-six and six three. So tall, he seems an unlikely rabbit, but the breadth of white face, the pallor of his blue irises, and a nervous flutter under his brief nose as he stabs a cigarette into his mouth partially explain the nickname, which was given to him when he too was a boy. He stands there thinking, the kids keep coming, they keep crowding you up.

His standing there makes the real boys feel strange. Eyeballs slide. They’re doing this for themselves, not as a show for some adult walking around town in a double-breasted cocoa suit. It seems funny to them, an adult walking up the alley at all. Where’s his car? The cigarette makes it more sinister still. Is this one of those going to offer them cigarettes or money to go out in back of the ice plant with him? They’ve heard of such things but are not too frightened; there are six of them and one of him.

The ball, rocketing off the crotch of the rim, leaps over the heads of six and lands at the feet of the one. He catches it on the short bounce with a quickness that startles them. As they stare hushed he sights squinting through blue clouds of weed smoke, a suddenly dark silhouette like a smokestack against the afternoon spring sky, setting his feet with care, wiggling the ball with nervousness in front of his chest, one widespread white hand on top of the ball and the other underneath, jiggling it patiently to get some adjustment in air itself. The cuticle moons on his fingernails are big. Then the ball seems to ride up the right lapel of his coat and comes off his shoulder as his knees dip down, and it appears the ball will miss because though he shot from an angle the ball is not going toward the backboard. It was not aimed there. It drops into the circle of the rim, whipping the net with a ladylike whisper. “Hey!” he shouts in pride.

“Luck,” one of the kids says.

“Skill,” he answers, and asks, “Hey, O.K. if I play?”

There is no response, just puzzled silly looks swapped. Rabbit takes off his coat, folds it nicely, and rests it on a clean ashcan lid. Behind him the dungarees begin to scuttle again. He goes into the scrimmaging thick of them for the ball, flips it from two weak grubby-knuckled child's hands, has it in his own. That old stretched-leather feeling makes his whole body go taut, gives his arms wings. It feels like he's reaching down through years to touch this tautness. His arms lift of their own and the rubber ball floats toward the basket from the top of his head. It feels so right he blinks when the ball drops short, and for a second wonders if it went through the hoop without riffling the net. He asks, "Hey whose side am I on?"'

Thus goes the beginning of John Updike's 'Rabbit' tetralogy: vivid, present, cinematic, magnanimous to multiple viewpoints. A sense of immediacy is bestowed upon the telling of *Rabbit, Run* (the title itself in the imperative mood initially captures the attention: run *where*? *why? at whose command?*) due to, in essence, the fact that the whole book is packaged in the present tense. Rabbit is, pretty literally, seen struggling to derive fulfillment in the contemporary 'adult' world - the 'cocoa suit' Rabbit is sporting, and his eagerness to 'fold it nicely' on the ashcan lid, of all places, and re/join the youngsters in a pickup scrimmage might as well symbolize a romanticized, half-remembered Eden of Rabbit's own former greatness on the court, scraping and snapping those stylish Keds, before a seeming 'Fall' into a moribund American dream. Comparisons have been rendered by literary critics and fans alike to Marcel Proust's *In Search of Lost Time* (a.k.a., *Remembrance of Things Past*) with Updike's early fiction in general and the 'Rabbit' tetralogy in particular; the early novels see Updike grappling with his own personal history and the 'Rabbit' novels increasingly function diachronically insofar as Rabbit's autobiography, the country's history, and Updike's reaction to both form an ongoing saga, fictionalized or otherwise, which forms the basis of further reflections and commentary.

Updike the man apparently suffered a crisis of faith while penning *Rabbit, Run* and found the theologian Karl Barth provided the necessary perspective to overcome what Updike then saw as the 'sense of horror that beneath the skin of bright and exquisitely sculpted phenomena, death waits.' The opening page of *Rabbit, Run* quoted above is preceded by an inscription from Pascal reading, 'the motions of Grace, the hardness of the heart; external circumstances,' which does, particularly in hindsight, presage the core spiritual dilemma at the novel's heart. According to the author, '*Rabbit, Run* was originally to be one of two novellas bound into a single volume; with its companion, *The Centaur*, it would illustrate the polarity between running and plodding, between the rabbit and the horse, between the life of instinctual gratification and that of dutiful self-sacrifice.' *The Centaur*, in turn, comes bookended with an inscription by Karl Barth about the centaur representing a creature formed as and responding to impulses from above and below. In a sense, the fact that Updike continued the 'Rabbit' tetralogy with three novels after *Rabbit, Run* is apt considering the country's trajectory in the second half of the twentieth century was more aligned with 'the life of instinctual gratification' than that of 'dutiful self-sacrifice.' The shift from a manufacturing to consumer economy is made much more explicit in the third and fourth installments of the Rabbit tetralogy, *Rabbit is Rich* and *Rabbit at Rest*, without Updike's being sententious: more of a fictional chronicle of what happened to the American Midwest or the 'Rust Belt.' What the literary critic Edmund Wilson said about Gustave Flaubert refraining from overt moral judgment very much applies to Updike's tetralogy of 'Rabbit' novels: 'It is not

true, as it sometimes supposed, that he [Flaubert] disclaimed any moral intention. He deliberately refrained in his novels from commenting on the action in his own characters: “The artist ought not to appear in his work any more than God in nature.” But, like God, he rules his universe through law; and the reader, from what he hears and sees, must infer the moral system.’ Updike himself literally says, writing from Georgetown nearly twenty years after *Rabbit, Run*’s original publication, ‘Rabbit is the hero of this novel, but is he a good man? The question is meant to lead to another - What is goodness?’ The reader, alone, is tasked with answering.

By the time the tetralogy’s second novel, *Rabbit Redux* (the title is inspired by an Anthony Trollope novel, *Phineas Redux*, as well as the John Dryden poem *Astraea Redux*), was published in 1971, the ‘irresponsible’ action depicted in *Rabbit, Run*, published in 1960, would seem fairly staid. Put another way, Rabbit now seems less a madman or callous libertine than a harbinger of the ‘if it feels good, do it’ permissiveness of the dropout 1960s counterculture - as the impish pied piper Timothy Leary memorably put it, ‘turn on, tune in, drop out.’ Can you dig? Updike could not quite entirely get behind the shenanigans and characters with whom the titular character, Rabbit, was now partaking. (Note the difference between what Updike says in a moment and the fact that these opinions do not match the characters’ feelings and, moreover, arrive in the form of a later preface, as opposed to the text proper.) Updike: ‘A number of people had asked me what happened to Rabbit, who was last seen running along a street of Brewer, Pennsylvania, in no special direction; now I would show them, and throw in all the oppressive, distressing, overstimulating developments of the most dissentious American decade since the Civil War - anti-war protests, black power and rhetoric, teach-ins, middle-class runaways [Rabbit takes in one, or arguably two, such ‘runaways’ in *Rabbit Redux*], drugs, and (proceeding eerily to its brilliant technology rendezvous through a turmoil of violence at home and aboard) the moonshot.’ The term ‘redux’ derives from the Latin ‘led back’: the reader is reminded, by Updike himself in the preface to *Rabbit Redux*, of the ‘hardness of the heart’ Pascal alluded to in *Rabbit, Run*’s opening inscription. The second installment of the Rabbit saga would appear the most political and angry of the lot. The good-natured, if arguably self-involved, hedonism, more widespread than ever, has curdled into an inability or unwillingness to be responsive to the needs of others, the larger community. Throughout *Rabbit Redux* Updike considers ‘the cost of the disruption of the social fabric’ and who pays the price. The situations and characters are not belittled or patronized; readers familiar with the era *will* recognize these archetypes.

In addition to the Pulitzer Prize awards for each, the third and fourth volumes in the ‘Rabbit’ tetralogy, *Rabbit is Rich* and *Rabbit at Rest*, were awarded the National Book Critics Circle Award for Fiction. Updike’s use of the present tense continues unabated, though the author now admits that the once daring technique has lost its ability to startle: ‘It is a delightfully apprehensive tense, quick on the pickup and easy to ride between external event and inner reflection’ and, furthermore, ‘Taking place in the present, they [the novels] are disturbingly open to accidents, to the random promptings of contemporary event, national and personal.’ Where has Rabbit ended up near the close of the twentieth century? After following in Earl Angstrom’s (Rabbit’s father’s) weary footsteps and becoming a linotypist in *Rabbit Redux*, Rabbit has apparently become ‘wealthy’, by Pennsylvania standards, in *Rabbit is Rich* by virtue of association with a local Toyota dealership falling into the hands of Rabbit’s reconciled wife,

Janice, whose father manned the tiller, at any admittedly shadier used-car lot, for decades before the Toyota partnership was inked. (The tetralogy, basically, follows four generations.) *Rabbit at Rest* sees Harry, easing further into an insouciant early retirement and logging countless hours on the links, splitting his time between Brewer and Deleon, a fictionalized locale in Florida, where Harry yearned to escape to decades prior in an impulsive break with an unsatisfactory homelife. (The reader will appreciate Rabbit's saga beginning and ending afoot a basketball court.) Updike meditating on Rabbit's trajectory: 'Using the materials and to some extent the maps and newspapers of our real world, fiction locates its characters in a cloudland where they can find the freedom to fulfill their tendencies. I wanted, in *Rabbit at Rest*, while plausibly portraying a specimen American male's evolutions into grandpaternity, frailty, lassitude, sensations of dispensability, and even inklings of selflessness, to allow the thematic tendencies, conscious or unconscious, of the three other novels to run to their destination, to wind up. I wanted to cap my series [in 1990] and make it a tetralogy while I still had most of my wits about me, and before my living connections with Pennsylvania quite ceased.'

The reader is somewhat amazed that the realist odyssey could be taken at all in the land of the free and home of the brave (or 'land of the dull and home of the literal,' as Gore Vidal was supposedly wont to cynically amend), let alone be engrossing across 1,500 pages, blessedly detached from the botched dialogue, bathos, stock characters, or longueurs one dreads in a tome of comparable girth. Why are these types of novels so difficult to pull off in the first place? John Updike, taking stock of the tetralogy a year after *Rabbit at Rest* was published, seemingly encapsulated part of the challenge in reflecting contemporary society in the late twentieth century: 'the very triumph of American popular culture in this century coats the individual, whom the writer would hold up to cultural study, in an impervious skin of cultural cliché; the crises of will and desire that occupy nineteenth-century novels are as it were kidded away by the image-saturated modern consciousness, which has pre-experienced everything, in a trivializing and dulling abundance of unearned sophisticated: this weariness of knowingness was wonderfully captured in the fiction of the late Donald Barthelme.' Updike's, like Flaubert's earlier, irony renders these narcotizing trends palatable. (Remember that even the mighty Herman Melville resorted to the high seas, after shoving off Nantucket, in order to achieve perspective on the United States - *Moby Dick* is frequently credited as being The Great American Novel, and there were apparently the same number of souls aboard the mighty Pequod as states in the union at the time of publication.) A more submerged, imponderable quality was also intuited by Updike: 'Something in our collective Puritan conscience insists, I fear, that to be worth writing our novels have to have, besides facts, an injection of the personal, the confessional, the spiritually urgent. It is this we grope for through our characters, this unprogrammable quality of *testimony* - of external evidences become interior signs - and to which a few characters do seem to grant us access.' The rabbit and the centaur appear to be well-chosen metaphors.

'Of all that might be omitted in thinking, the worst was to omit your being.' -Saul Bellow, *What Kind of Day Did You Have?*

'Art and religion first; then philosophy; lastly science. That is the order of the great subjects of life, that's their order of importance.' -Muriel Spark, *The Prime of Miss Jean Brodie*

Unidentified Anomalous Geese (UAG)
… and other conspiracy theories

The ontology of the universe is not that of a simulation,
but a simulation of a simulation of a simulation . . .
— of an explanation
— just a Conspiracy Theory.

“Truth is the daughter of time” or is it crime, sometimes?
If there were excess deaths,
among young safe and effective working people,
a vaccine for statistics would be needed.

If Lockheed Skunkworks (“Lockheed Martian”)
is rumored to have reverse engineered Unidentified Anomalous Geese (UAG),
often said to be much larger on the inside than on the outside,
follow the missing money.

Supposedly someone not called dullest
suggested the existence of Unidentified Anomalous Geese (UAG)
should be called just a “conspiracy theory.”
Hence, a conspiracy theory for the origin of “conspiracy theory.”

In the conspiracy of no conspiracy,
Oswald, acting alone,
fired from multiple locations,
by quantum superposition.

The “I” *is* a simulation of a simulation of a simulation . . .
Each breath may become a digital ID.
Tat tvam asi?

May-Tzu

‘Ecstatic bird songs pound
the hollow vastness of the sky
with metallic clinkings—
beating color up into it
at a far edge,—beating it, beating it
with rising, triumphant ardor,—
stirring it into warmth,
quickening in it a spreading change,—
bursting wildly against it as
dividing the horizon, a heavy sun
lifts himself—is lifted—
bit by bit above the edge
of things,—runs free at last
out into the open—!lumbering
glorified in full release upward—
songs cease.’

-William Carlos Williams (‘Dawn’)