

Improving the Gallup World Poll

By David Blanchflower

See page 15 to find the Appendix.

Moving forward, I am hoping to calm this debate and focus on science and what we can learn about the human condition. My post on the Gallup World Poll (GWP) may have been overly confrontational, so this attempt addresses the critics constructively and, I hope, provides a way forward. The title of my last post, "*Is The Gallup World Poll Reliable?*" was also too provocative, and my long-time colleague Alex told me off for it! My goal is to engage constructively and foster dialogue to improve the Gallup World Poll, which in many respects, is a pathbreaking survey. My goal is to help to make it better in a post-Covid world.

In my original post, I also critiqued the *misuse* of GWP data in arguments that downplay concerns about the impact of digital technology on youth mental health. Rothwell and Srinivasan did not disagree about the misuse of the GWP — they disagreed about my broader claims around the reliability of the poll.

I do not intend to critique Gallup data in a way that implies it cannot shed light on important social and economic issues: it can. The Gallup World Poll is an invaluable data source for research, which I have utilized many times. Like any data file with extensive coverage across countries, it has strengths and weaknesses. The GWP's greatest strength is its broad coverage, especially in less developed regions where data is scarce, allowing researchers to track changes over time.

However, when a survey like the GWP covers numerous countries with varying populations using the same sample size, it becomes challenging to make detailed and accurate conclusions about specific subgroups, such as age groups or genders, due to insufficient sample sizes. As Gallup researchers Jonathan Rothwell and Rajesh Srinivasan point out, the individual country files are not intended to provide representative estimates by state, month, age, or gender.

Another significant limitation is the nature of the survey questions, most of which are just on/off variables. First, the well-being variables are not well-suited for a world that has experienced four major shocks: the Great Recession, the digital revolution, COVID-19, and a European war. Second, the survey's ability to track the declining mental health or unhappiness of young people is inadequate because they include no variables that can identify those struggling with serious mental health problems, who potentially are at risk of self-harm or even homelessness. The GWP struggles to identify

negative effects, especially when severe mental health symptoms are of particular interest. We don't know why yet, but it is harder to detect severe symptoms using questions about happiness/positive affect than with unhappiness/negative affect data. Even though the GWP has questions on pain, stress, sadness, and worry, they are not adequate.

The Cantril variable, a heavily used positive affect measure from the GWP, assesses well-being as a "ladder of life" but is not well-suited to capture the broad changes since 2008. It asks how someone's life has turned out so far rather than something more immediate, which is impacted by life events like unemployment or financial worries. Indeed, the variables are often contradictory. Of the 426 respondents to the US GWP from 2015-2023 who reported a Cantril score of 3 or less on a scale of 0-10, 206 (48%) said they had enjoyment yesterday. Of the 4,275 who scored 8 or higher on Cantril 1073 (25%) said they experienced worry yesterday. These kinds of contradictions are unusual in well-being surveys.

The four main positive affect variables - enjoyment, smiling and laughing, well-rested, and happiness - and the four negative affect variables - worry, sadness, stress, anger, and pain - are all binary yes/no (1,0). Moving forward, this can be improved by including new measures of ill-being, such as the GHQ score used in the UK panel Understanding Society and the English and Scottish Health Surveys; the MHQ score used in Global Minds; the Kessler Score used in the NHIS and NSDUH; and the PHQ and GAD scores used in Come-Here, which are commonly used validated screening tools.¹ If not, why not?

It seems the well documented U-shape in happiness and the hump shape in unhappiness in age have now gone. We have found evidence that in the years since 2020 especially, that ill-being now declines with age in 90 countries and the young are especially unhappy.

I do need, though, to respond to several of the points made by the Gallup team.

Point-by-point responses to Jonathan Rothwell, and Rajesh Srinivasan

My major claim was that the GWP seems to be an outlier once I moved to an analysis of youth well-being, which does not seem to be picked up in the GWP. I argued there were four reasons:

1. Samples are small, especially for large countries, and do not appear to be nationally representative.

¹ <https://www.aafp.org/pubs/afp/issues/2018/1015/p508.html>

2. The survey's well-being measures did not decline during the Great Recession or COVID-19, which seems unlikely to reflect reality, given that most other surveys show declines.²
3. The timing of the survey varies by month within and between countries.
4. The GWP shows that internet "access" is positively correlated with well-being.

Rothwell and Srinivasan responded to these criticisms, and I will address their main points in turn.

#1. Academic Publications: *"We know he has frequently published academic articles using Gallup's global wellbeing data, including one this year, and has evidently not retracted any of them."*

This is true. In a [previous paper](#), for example, my co-authors and I showed that GWP data for 2008–2017 found U-shape curves in happiness across the lifespan, and the results, as reported, are correct. However, we have also pointed out flaws with the Gallup World Poll in several papers, including its [lack of responsiveness to shocks](#) and [issues with gender effects](#) and [well-being rankings](#). Once I became aware of these problems, I stopped using GWP and started using other sources like ISSP, Global Minds, and Come-Here, which do not appear to have the same issues. These data files show that the U-shape has gone, less so in the GWP.

#2. Sample Size: *"The World Poll annually conducts roughly 1,000 interviews with a nationally representative sample of adults per country, per wave. This is more than enough to provide population-wide estimates at the national level, which is what the World Poll is designed to measure"*

It is the case that a well-designed survey can allow researchers to produce population estimates with 1,000 unweighted respondents, provided we are able to relate the achieved sample to the population through sample probability weighting and adjustments for non-response. I have been concerned that many using GWP are not as aware as they should be of the survey's design and the methodology used to reweight the raw data with the weights provided. I'm keen to see more information about how successful the reweighting strategy is in allowing us to extrapolate from the sample to country populations, and to know more about adjustments for survey non-response.

² Plus, these two shocks were the biggest negative shocks in a generation, and it is hard to believe they didn't lower wellbeing. We know, for example that a rise in the unemployment or inflation rates has a major negative impact on wellbeing – see Blanchflower DG, DNF Bell, A Montagnoli, and M Moro, '[The happiness tradeoff between unemployment and inflation](#)', *Journal of Money Credit and Banking*, Supplement to vol 46(2), October, 2014, pp. 117-141

It is certainly true that we researchers are always greedy for more. But sometimes there is a good reason. In this case, as Rothwell and Srinivasan say, the precision of estimates is dependent upon sample sizes, and a survey's ability to pick up statistically significant changes of a certain size is dependent upon sufficient unweighted sample sizes. So when we think of a sample size being 'more than enough' we need to remember that it might not pick up significant changes when those changes are imprecisely estimated. While 1,000 interviews per country may be sufficient for population-wide estimates at the national level, small sample sizes can lead to concerns about the stability of estimates, the size of coefficients, and the ability to generate disaggregated estimates.

For example, the Current Population Survey, produced by the Bureau of Labour Statistics and used to calculate the monthly unemployment rate, in 2023, for example, had *an annual* sample of approximately 1 million respondents aged 16 and over, a thousand times bigger than the US GWP. This survey, which is publicly [available for researchers at no charge, is sufficiently](#) at no charge, does allow disaggregated monthly and state-level estimates as well as by age and gender, and race and ethnicity to be published and analyzed.

The other issue, of course, is being able to examine well-being and other outcomes of interest in sub-populations. This was more easily done with, the US Daily Tracker conducted by Gallup, which had 350,000 observations a year, far more than the 1,000 for the US in the World Poll. The global ranking of the US was very different when we used the data from the Daily Tracker than it was when using USGWP data on all of the GWP wellbeing variables. In our Wellbeing Rankings paper, we found systematic differences in the US scores between those reported in the GWP and the Daily Tracker, with those in the Daily Tracker reporting much lower negative affect. (In all cases, the [differences between](#) the US scores in the two surveys are statistically significant).

#3. Youth Well-Being: “To test whether a 16-percentage-point decline in the share of youth who are thriving in wellbeing is significant in the United States, one only needs 143 respondents per period. The World Poll’s U.S. sample includes 470 people between the ages of 15 and 24 from 2018 to 2023 and 561 between 2005 and 2011.”

Rothwell and Srinivasan do not deny the sample sizes for youth in the US GWP are small. From 2018 to 2023, there were only 470 people aged 15-24 in the US GWP sample, averaging 78 per year. A larger sample size would provide more reliable insights, especially for phenomena impacting specific subgroups like less-educated females. For instance, the GWP has three education categories: completed elementary education or less, completed secondary education, and completed four years of higher

education. In the US in 2020, zero females under 25 had not completed elementary education, one in 2021 and 2023, and two in 2022. One also needs to be concerned not just about significance but also about coefficient size and stability, which are much bigger issues in small samples than bigger ones.

Rothwell and Srinivasan say that because of small sample sizes, the World Poll 'is not designed for state-level analysis,' or indeed monthly analysis, which presumably means it is not designed for analyzing subgroups, including the young. This also means it is not designed for analysis conducted on very small cells using these data, [as done by some researchers](#).

I believe I am correct in saying that GWP does not stratify its sampling by location eg. by state in the case of the USA. Consequently, the small sample sizes means one cannot undertake analyses of geographical variance in wellbeing. That's a shame because we know for economic and social reasons they can be important. The way around this would be either to raise sample sizes very substantially so that one can think of the sample as 'as good as random' or else stratify by location such that representativeness at local level will be assured. But that would only be meaningful if the sample size was large enough to be sure that one had a sizeable N of respondents in each state cell.

#4. Comparison with World Values Survey (WVS): “*The World Values Survey (WVS) averaged 1,705 responses per country for Wave 7, collected over six years from 2017 to 2022 in 92 countries. During that same period, the World Poll collected about 6,000 responses per country across 142 countries. Blanchflower has not mentioned any concerns about the WVS sample size in his recent publications, despite its much smaller size than the World Poll.*”

Rothwell and Srinivasan pointed out that the WVS averages 1,705 responses per country for Wave 7. It is true that [in a previous paper](#), my co-authors and I used the WVS data to look at U-shapes between 1990-2014. However, the structure of the WVS is different from the GWP; each country was only interviewed in one of the years during the wave period. For example, in the Great Recession years, the US was interviewed in 2006 and not again until 2011, missing significant periods. This is why [we have not used](#) the WVS to analyze responses to shocks.

#5: Inconsistency with the Gallup US Daily Tracker

Rothwell and Srinivasan also do not address my point from my previous post (also [here](#)) that the findings in the US GWP are entirely different from those in the much larger Gallup US Daily Tracker that have the same questions. That is a concern, especially as

we have no other large surveys for any other country to check results. This is likely a major issue in large countries that have the same sample size as small countries – the US has approximately the same sample size as Malta and Cyprus.

#6. Use of a Single Well-Being Measure: “*Blanchflower’s use of a single measure of wellbeing — rather than two (current and future) — further injects error.*”

Rothwell and Srinivasan claimed that using a single measure of well-being injects error. It is unclear what error this introduces. In prior work, we have used various measures, including Cantril, life satisfaction, and positive and negative affect variables. We found that different variables produce different rankings, highlighting inconsistencies within the GWP. The WHR report uses single measures. In the expansive literature of over 625 published papers on happiness U-shapes I don’t believe any used thriving. (cite Blanchflower Graham and Piper, 2023).³

I recommend this approach because relying on single metrics can be dangerous, as Rothwell and Srinivasan says. Wellbeing is multifaceted, and we might reasonably expect rankings to vary quite a bit across metrics. This is what Ed Diener taught us many years ago. So, when we study wellbeing, we like to use as many of the available metrics as possible to fully understand what’s going on.

To illustrate, in [our analysis from 2008-2017](#), Finland ranked 4th on the Cantril measure, 93rd on enjoyment, 99th on laughing and smiling, 122nd on being well-rested, 51st on pain, 15th on sadness, 130th on worry, 1st on anger, and 51st overall, where we rank the negative affect variables from least to worst.

When we looked at the GWP data from 2021-2023, Finland ranked 1st and Denmark 2nd using the Cantril measure out of 145 countries. However, when we examined smiling and laughing, Finland ranked 73rd and Denmark 55th. If we used worry, ranking the least worried as top-ranked, Finland ranked 43rd and Denmark 41st. Alternatively, using sadness or stress produced different rankings. For example, in the smiling and laughing measure, Guatemala and Panama were the top two, and Kazakhstan and Taiwan were the least worried countries. The various GWP measures are inconsistent with each other which is a problem

An example of another inconsistency comes from Eurostat’s data. In the third wave of the European Health Interview Survey, 7.2% of the European population reported experiencing chronic depression, with a higher proportion of women (8.7%) than men (5.7%). Country rankings showed Iceland with the highest incidence of chronic depression among 31 countries, followed by Portugal and Sweden, with Romania

³ Also Vuorre and Przybylski only used single measures of wellbeing - <https://tmb.apaopen.org/pub/a2exdggg/release/1>

having the lowest incidence. In contrast, the GWP rankings in their 2019 report showed Sweden (7), Iceland (4), Denmark (2), and Finland (1) ranked very high for overall well-being, with the rankings from the 2020 World Happiness Report using GWP data for 2017-2019 in parentheses.⁴ Romania ranked 48th. Despite claims that Finland is the happiest country in the world based on GWP data it doesn't seem to be. It also has high suicide rates.

#7 “The first point is bizarre because it is so easily refuted by publicly available summary data for the United States. According to World Poll data, the share of U.S. adults thriving in well-being fell from 66% in 2007 to 61% in 2008, just after the Great Recession began, and dipped further to 57% in 2009, when unemployment spiked.”

The authors say this is ‘easily refuted’ but do not refute anything that I stated, as I did not make any assertion about ‘thriving’ in my post. Indeed, I have never used thriving in any of my thirty peer-reviewed papers and know of no other peer-reviewed papers in the economics of happiness that use that concept. I was focusing on the lack of change in other single variables, not two combined unvalidated variables, in the last post.

#8. Thriving and Other Measures: “According to World Poll data, the share of U.S. adults thriving in wellbeing fell from 66% in 2007 to 61% in 2008, just after the Great Recession began, and dipped further to 57% in 2009, when unemployment spiked. While it is true that the World Poll data for the United States do not show a drop from 2019 to 2020 or 2021 in overall wellbeing (percentage thriving), World Poll data collected in May 2019 and May 2020 in the United States show a large increase in the share of adults who “worried a lot” the day before the survey; increasing from 35% to 43%.”

My colleagues and I have not used the 'thriving' variable before. Nor is it widely used; indeed, it isn't included at all in the World Happiness Report.⁵ Of course, I stand to be corrected. It is indeed true that both thriving and Cantril rose for the US in the pandemic, as the authors claim in their note, and we confirm that worry did rise. But it turns out that stress and pain both fell.

Here is an interesting observation for Europe, including both Western and Eastern Europe, when weights (variable =wgt) are applied. Both Cantril and thriving scores rose

⁴ <https://worldhappiness.report/ed/2019/>

⁵ Thriving is defined as 1 if Cantril ≥ 7 and if life in five years is 8 or higher, zero otherwise, why? The basis for using and calculating this measure though seems somewhat arbitrary. See for example this by Shelton and Herrick, in relation to the GHQ <https://www.sciencedirect.com/science/article/abs/pii/S0033350609002674>

during the Great Recession and the COVID-19 pandemic, which seems unlikely given the economic and social disruptions caused by these events.

	Cantril	Thriving
2007	5.86	.37
2008	6.91	.67
2019	6.50	.44
2020	6.53	.46

Table. Cantril and Thriving throughout Europe during the Great Recession and COVID.

#9. Month-to-Month Analysis: *“As non-World Poll Gallup data for the United States show, wellbeing fluctuated greatly from month to month and even day to day, according to things like spikes in unemployment, hospitalization counts, pandemic relief checks, business re-openings, and vaccination announcements and implementation. It would be foolish to expect a point-in-time national survey to reflect all of this variation, and Gallup never intended for the World Poll to be used to analyze monthly national trends.”*

One benefit of large sample surveys is that they can be used to analyze month-to-month variations, as we have [done](#) in [previous studies](#) on the gender well-being gap. The Eurobarometer, for example, allows for month-by-month analysis. In [one paper](#), we show major inconsistencies in the GWP overall, where we find a positive female differential using happiness and worry measures but a negative one using enjoyment measures.

I think good progress could be made in reconciling some of the inconsistencies in results across surveys if the timing of fieldwork, and the specific time of each survey interview, were available in the data. That way the researcher is about to account for timing differences in fieldwork when analyzing the data. But it would also help if survey fieldwork could be coordinated more: other surveys such as the Eurobarometers accomplish that.

#10. Mental Health Trends: *“Blanchflower states that, ‘The decline in the well-being of the young means that mental health now improves with age.’”*

This is correct, as noted in my first post. Alex Bryson, other co-authors, and I have found this to be true in multiple datasets across 90 countries since 2020. This contrasts with what we found in an earlier paper (see [Blanchflower, D.G., Graham, C.L. The Mid-Life Dip in Well-Being: a Critique. Soc Indic Res 161, 287–344 \(2022\)](#)), which found U-shapes in well-being across 155 countries using Cantril and stress from the GWP, 2005-2019 - their Table 3. As we show below, we now also find U-shapes in over a hundred countries in the GWP from 2015-2023 using thriving.

In other datasets, we [find evidence](#) that ill-being declines with age while well-being rises. In [one paper](#), we find that in the International Social Survey Programme series pooled country files, the percentage of people saying they were often or very often unhappy and depressed in the prior four weeks rose with age in 2011 but fell with age in 2021. This change is illustrated in Chart 1, clearly indicating a shift over time.

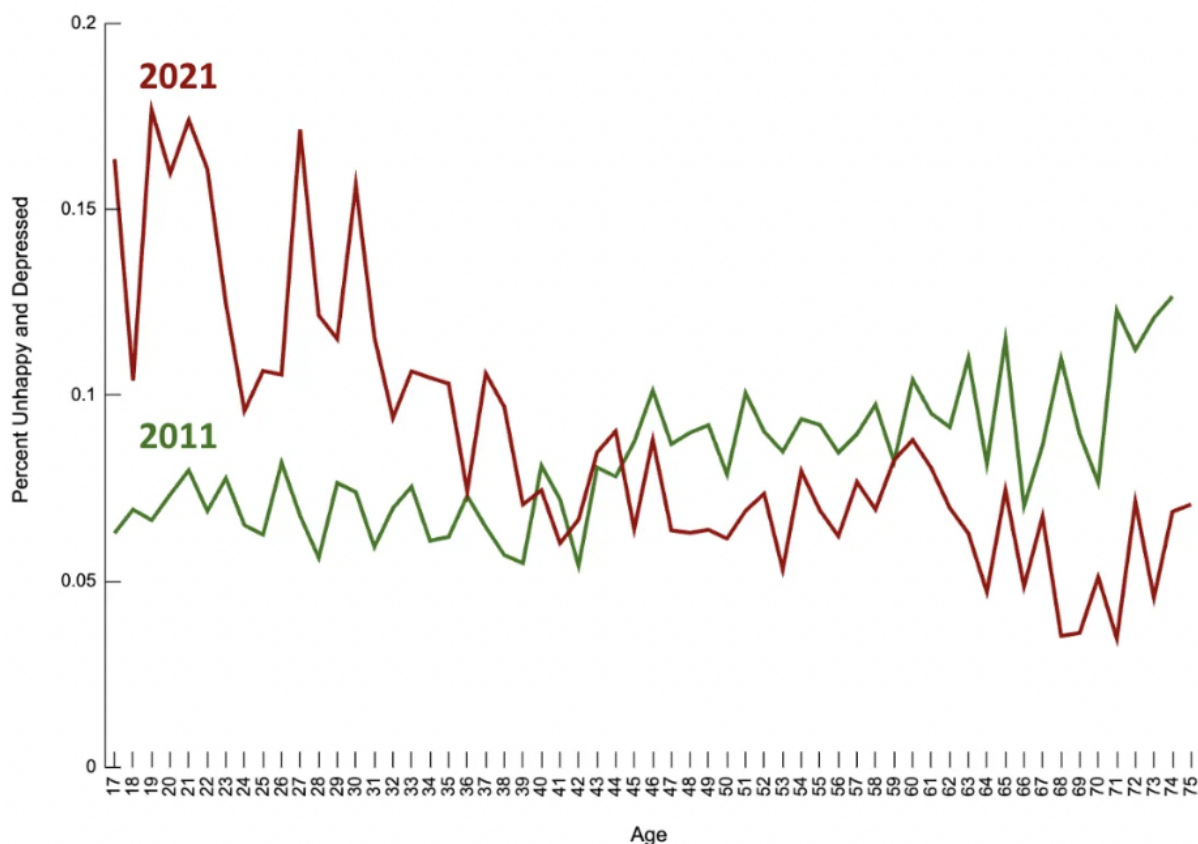


Chart 1. Percent change in those who report being “often and very often unhappy and depressed” by age (2011 and 2021). Source: International Social Survey Programme.

Declining well-being among youth means that mental health now improves with age, contrasting with previous findings of a U-shape in happiness. Various datasets show this trend, indicating significant changes over time.

#11. PISA Data Comparison: “To assess whether PISA data align with World Poll data, we downloaded the latest PISA data (2022) and created a variable “highly satisfied with life” for 15-year-olds who responded with a 7 or higher on a zero to 10 scale when asked by PISA’.

I also downloaded the 2022 PISA data and the 2018 data, which contained the life satisfaction data that Rothwell and Srinivasan examined. The PISA data over these years, for fifteen year olds showed big declines (the size of the effect matters!) in life satisfaction among 15-year-olds in 55 out of 62 countries, as reported in Appendix 1. This aligns with our findings of declining youth well-being, which the GWP fails to capture adequately.

#12. Thriving Regression Analysis

I have not used the 'thriving' variable before, and to my knowledge it is not used in the World Happiness Report, or by any economists, for example, but since the authors consider it their go-to measure, let's examine it. To determine if the GWP shows declining well-being among the young, we ran a regression analysis with thriving as the dependent variable using the GWP from 2015 in each country, with thriving regressed on age, age squared, gender, and year. Using age and age squared is the standard way to test for U-shapes with the expected signs negative for age and positive for age squared with a minimum around age fifty.

We found that for the thriving variable that there are U-shape in age, since 2015 in 112 countries – meaning that the age term was significantly negative and the squared term significantly positive with t-values > 1.65. In 46 countries, there was no significant U-shape relationship.⁶ This includes nearly all of the major advanced countries we studied in the European Come-Here file, such as France, Germany, Italy, Spain, and

⁶ The 112 countries that we found U-shapes for are Afghanistan; Albania; Algeria; Argentina; Armenia; Azerbaijan; Bahrain; Bangladesh; Belarus; Benin; Bolivia; Bosnia Herz; Botswana; Brazil; Bulgaria; Burkina Faso; Cambodia; Cameroon; CAR; Chad; Chile; China; Colombia; Comoros; Congo Brazzaville; Costa Rica; Croatia; Cyprus; Dominica; Ecuador; Egypt; El Salvador; Estonia; Eswatini; Ethiopia; Gabon; Georgia; Greece; Guatemala; Haiti; Honduras; Hungary; India; Indonesia; Iran; Iraq; Ivory Coast; Jamaica; Japan; Jordan; Kazakhstan; Kenya; Kosovo; Kuwait; Kyrgyzstan; Laos; Lebanon; Lesotho; Liberia; Libya; Madagascar; Malawi; Malaysia; Mali; Malta; Mauritius; Moldova; Mongolia; Montenegro; Morocco; Mozambique; Myanmar; Namibia; Nepal; Nicaragua; Nigeria; North Macedonia; Pakistan; Palestine; Panama; Paraguay; Peru; Philippines; Portugal; Puerto Rico; Romania; Russia; Rwanda; Senegal; Serbia; Sierra Leone; Slovenia; Somalia; Sri Lanka; Taiwan; Tajikistan; Tanzania; Thailand; The Gambia; Togo; Tunisia; Turkey; UAE; Uganda; Ukraine; Uruguay; Uzbekistan; Venezuela; Vietnam; Yemen; Zambia and Zimbabwe.

Sweden, plus the US, UK, Canada, Australia, New Zealand and the Netherlands whose national agencies have reported declining ill-being in age.⁷ It is notable also that in contrast for the majority of those that have U-shapes we find using Global Minds and ISSP, BRFSS and UK Understanding Society data no evidence of U-shapes in age but declining ill-being in age and rising well-being in age throughout the lifespan. The GWP is an outlier.

#13. Internet Use and Disaggregated Estimates:

The GWP's suitability for disaggregated analysis is questionable, particularly for country/year/age/gender cells. Other datasets, like Global Minds, offer more comprehensive data, including a variety of mental health measures. This issue is *less about the Gallup World Poll Itself and more about [the way researchers have misused it](#)*.

#14. Global Analysis: “*Blanchflower does not even attempt to test the reliability of the World Poll in the 141 or more countries with data outside the United States.*”

In fact, I've co-authored a number of papers which compare GWP and other data across multiple countries. That is precisely what Alex Bryson and I have done in a number of earlier papers, especially [here](#) and [here](#). A major weakness of the GWP is the very poor quality of questions used, particularly for negative affect, which are binary (1,0) variables.

The issue with questions like 'Did you experience enjoyment yesterday Y/N' is that they are not sensitive to changes in the degree of enjoyment. Almost everyone experiences some enjoyment every day, so the question doesn't capture shifts in well-being. This makes the data not unreliable but not very useful.

The problem with Yes/No variables is that they miss changes at the extremes, such as severe mental health issues. For example, in the 2022 BRFSS, 37% of those under 25 said they had no bad mental health days in the past month, but 8.9% said all 30 days were bad, up from 3.5% in the 1993-2000 period.

In the Global Minds surveys from 2020-2024 across more than 100 countries, mental health is measured on a 1-9 scale, with an overall score ranging from -100 to +200. Similarly, the NHIS 1997-2018 surveys use the Kessler-6 measure, which ranges from 0-24, to identify serious mental distress. The National Survey on Drug Use and Health

⁷ The 46 countries that we did not find u-shapes for are Australia; Austria; Belgium; Bhutan; Burundi; Canada; Congo Kinshasa; Czechia; Denmark; Finland; France; Germany; Ghana; Guinea; HK; Iceland; Ireland; Israel; Italy; Latvia; Lithuania; Luxembourg; Maldives; Mauritania; Mexico; Netherlands; New Zealand; Niger; Northern Cyprus; Norway; Poland; Qatar; Saudi; Singapore; Slovakia; South Africa; South Korea; South Sudan; Spain; Sweden; Switzerland; Syria; Trinidad; Turkmenistan; UK and USA

(NSDUH), as reported by [Jean Twenge and co-authors](#), also utilizes similar measures. The Come-Here survey uses the PHQ-9 (0-27) for depression and the GAD-7 (0-21) for anxiety. The UK Household Longitudinal Survey uses the GHQ score (0-36), with scores of 20 or higher indicating 'despair'. In 2009, 8% were classified as in despair, compared to 12% in 2021. The GWP does not capture these changes well, especially among the most vulnerable.

As I said before, the Gallup World Poll is problematic for understanding global ill-being, especially since 2014. This is especially true in less developed countries. The weakness is especially apparent in documenting the worsening mental health of the young, where the GWP seems to be the outlier. Other data sources show that the U-shape in happiness and the hump shape in unhappiness have disappeared, with happiness now increasing with age. The GWP has largely missed this trend.

Proposals for Improvement

Having been an avid user of GWP to rank countries by wellbeing; to examine the age and sex profiles of wellbeing over time; and considered the links between wellbeing and behaviors such as internet usage, both in terms of intensity and content. I have some proposals which might, if implemented, increase the value we obtain from GWP data.

- Consider raising sample sizes for GWP countries and, in the meantime, provide clear guidance to users on how GWP is designed and thus what the analyst can and can't do when studying the role of age and location.
- Revisit both the nature of the wellbeing questions in GWP and the way that wellbeing questions are coded and validated. Over-reliance on (0,1) outcomes and over-reliance on Cantril's ladder mean we obtain only a partial view on what's going on in the world in terms of people's well and ill-being. Possibilities are to include questions like MHQ, PHQ, GHQ, GAD which will permit comparison with the findings from the existing variables.
- I would experiment with different positive affect questions including life satisfaction and happiness on 1-10 scales perhaps also experimenting with

time periods covered.⁸ This is likely preferable than asking about the integral of satisfaction of the life course.

- Seek to align the timing of GWP surveys and, if this is not possible, issue guidance on the timing of survey fieldwork to researchers so they can incorporate this information into their analyses.
- Introduce dedicated questions to address new challenges such as those relating to internet usage.

Conclusion

The trends or declining well-being of the young seem to have started around 2015 or so, and hence was not caused by, but was likely exacerbated, by Covid.

All of the datasets, researchers have used around the world to document well-being in the post-COVID world, have strengths and weaknesses. Some have data available prior to Covid but many don't e.g. Global Minds and the US Household Pulse Surveys:

My research strategy over the years has been to look at multiple data sources. When estimates are contradictory, researchers need to collaborate to understand why. In this instance, the GWP seems to be an outlier, and we need to understand the reason for that.

I am happy to engage in a constructive dialogue about appropriate questions to use, the form they should take (not 1,0 binary yes/no questions), what weighting schemes to use to make them representative (which I am currently working on), and what level of disaggregation is appropriate. We also need to understand the extent to which findings on positive affect and happiness are consistent with those on negative affect, including unhappiness.

There are more serious issues at stake than a disagreement over one Gallup poll. The mental health of adolescents is a grave concern, with youth depression and suicide rates having more than doubled. Addressing these critical issues requires accurate and reliable data, and improving our understanding and measurement of well-being is a crucial step toward better mental health outcomes for our youth in particular.

⁸ For example, in the UK the national Annual Population Survey conducted by the Office of National Statistics the life satisfaction question says *Overall, how satisfied are you with your life nowadays?* While the happiness question says *Overall, how happy did you feel yesterday?*

<https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/methodologies/personalwellbeingintheukqmi>

My intention is to help improve the GWP, not to tear it down. We are all in this together.

Appendix 1

Life satisfaction data in 62 countries, 2022 and 2018 PISA data.

	2022	2018	Change
Albania	8.01	8.61	-.60
Argentina	6.69	7.26	-.57
Austria	6.69	7.14	-.44
Azerbaijan	6.80	7.24	-.44
Brazil	6.85	7.05	-.21
Brunei D	5.86	5.80	+.06
Bulgaria	7.04	7.15	-.11
Chile	6.41	7.03	-.62
Chinese Taipei	6.85	6.52	+.33
Colombia	6.96	7.62	-.66
Costa Rica	7.32	7.96	-.64
Croatia	7.37	7.69	-.32
Czechia	6.56	6.91	-.36
Dominican R	7.44	8.09	-.65
Estonia	6.91	7.19	-.28
Finland	7.41	7.61	-.21
France	6.77	7.19	-.42
Georgia	7.62	7.60	+.02
Germany	6.51	7.02	-.51
Greece	6.62	6.99	-.37
Hong Kong	6.49	6.27	+.22
Hungary	7.21	7.12	+.10
Iceland	6.90	7.34	-.44
Indonesia	7.22	7.47	-.25
Ireland	6.59	6.74	-.15
Italy	6.53	6.91	-.38
Japan	6.76	6.18	+.58
Jordan	6.77	6.88	-.11
Kazakhstan	8.41	8.76	-.34
Korea	6.36	6.52	-.16
Kosovo	7.87	8.30	-.44
Latvia	6.76	7.16	-.40
Lithuania	7.14	7.61	-.47
Macao (China)	6.41	6.07	+.34

Malaysia	6.63	7.04	-.40
Malta	6.24	6.56	-.33
Mexico	7.26	8.11	-.85
Montenegro	7.52	7.69	-.16
Morocco	6.76	6.95	-.19
Netherlands	7.29	7.50	-.21
North Mac	7.65	8.16	-.51
Panama	7.04	7.92	-.88
Peru	6.37	7.31	-.94
Philippines	6.97	7.21	-.24
Poland	6.26	6.74	-.49
Portugal	7.06	7.13	-.07
Qatar	6.77	6.84	-.07
Moldova	7.01	7.68	-.67
Romania	7.53	7.87	-.34
Saudi Arabia	7.36	7.95	-.58
Serbia	7.48	7.61	-.13
Slovakia	7.02	7.22	-.20
Slovenia	6.61	6.86	-.25
Spain	6.88	7.96	-1.07
Sweden	6.91	7.01	-.11
Switzerland	7.06	7.38	-.31
Thailand	7.12	7.64	-.51
Türkiye	4.90	5.62	-.72
Ukrainian	7.35	7.54	-.19
UAE	6.85	6.88	-.03
UK	6.07	6.16	-.09
Uruguay	7.03	7.54	-.50