

NBER WORKING PAPER SERIES

AN EVALUATION OF BIASES IN WELLBEING ESTIMATES USING INTERVIEWERS
VERSUS ONLINE DATA COLLECTION IN THE GLOBAL FLOURISHING STUDY

David G. Blanchflower
Alex Bryson
Alan J. Cui

Working Paper 34599
<http://www.nber.org/papers/w34599>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2025

We thank Chris Felton, Noah Padgett, Jon Skinner, Doug Staiger and Tyler VanderWeele for helpful comments and suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2025 by David G. Blanchflower, Alex Bryson, and Alan J. Cui. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

An Evaluation of Biases in Wellbeing Estimates Using Interviewers Versus Online Data Collection
in the Global Flourishing Study

David G. Blanchflower, Alex Bryson, and Alan J. Cui

NBER Working Paper No. 34599

December 2025

JEL No. I3, I30

ABSTRACT

Survey mode is a critical consideration in well-being research using the 22 country Global Flourishing Study (GFS). We find marked differences in responses to well-being questions if they are obtained via a telephone interviewer (CATI) or online (CAWI). In fifteen countries in the GFS both survey modes are used. In seven of the most advanced countries, only CAWI is used. Well-being responses are markedly different across survey modes within countries using both modes and compared to countries only using CAWI. On average, CATI tends to produce higher well-being scores. Failure to account for these sampling differences biases results. This is the case in twenty-three prior studies that used the GFS and took no account of survey mode. We examine each of these studies and show the findings differ by survey mode. Combined with the presence of other unobservable country-level confounders, the differences between survey modes substantively weakens the internal validity of cross-country comparisons and random effects meta-analysis conducted with the GFS.

David G. Blanchflower
Dartmouth College
Department of Economics
and Adam Smith School of Business,
University of Glasgow
and also NBER
blanchflower@dartmouth.edu

Alan J. Cui
Dartmouth College
alan.j.cui.27@dartmouth.edu

Alex Bryson
University College London
a.bryson@ucl.ac.uk

1. Introduction

It is common to use a range of survey methodologies to elicit information from survey respondents in household surveys. Traditionally survey agencies sent questionnaires through the post in the hope sampled individuals would complete the survey and return it in a pre-paid envelope, or else they sent survey interviewers to sampled individuals' homes to conduct the survey face-to-face, recording responses on paper. With the advent of new computer-based technologies agencies began to shift to computer-assisted interviews, where responses were recorded on a laptop so that data could be more easily uploaded to a computer database. Further technological innovations made it cheaper for agencies to collect data over the telephone (computer-assisted telephone interviews, CATI). Most recently, with the rollout of fast-speed broadband connections most populations can be sampled and contacted on the internet, facilitating Computer-Assisted-Web-Interviews, or CAWI. The latter entails considerable savings for agencies and the entities employing them, thus incentivizing the 'push to web' whereby sampled individuals are encouraged to respond via the web, only resorting to other means where this is infeasible. If a web-based approach does not elicit responses, agencies may resort to CATI.

Whilst cost-savings are available via CAWI when compared to other methods, the choice of survey mode can have consequences for the quality and nature of data collected. Survey response rates may vary across CATI and CAWI, for example, and may vary in different ways across sub-populations. For instance, the age profile of respondents may vary if older people are less comfortable using the internet, or if young people are less likely to answer the phone and more likely to engage with the internet. Conditional on responding, it is also possible that the nature of individuals' responses will depend on how they are being asked. Responses may be more prone to social desirability bias if conducted in the presence of a survey interviewer, compared with scenarios in which no interviewer is present (Rickwood and Coleman-Rose, 2023).

Recent studies have shown that the ways in which individuals report their mental health and subjective well-being is sensitive to survey mode. For example, Blanchflower and Bryson (2025a) found that, for a sample of individuals in the Global Flourishing Survey (GFS) in the Middle East and Asia, the young aged 18-24 were less depressed and happier than those aged 25 and above with CATI, but the reverse was the case when the survey was conducted via CAWI. This sensitivity was not apparent with respect to other variables such as smoking, drinking, pain and exercise.¹

In this paper, we focus on the role of survey mode to improve understanding of variance in well-being within and across countries in the Global Flourishing Study (GFS). In the majority of countries in the GFS, which are mostly emerging economies, both survey modes are used. In seven of the most advanced countries, only CAWI is used. Results are considerably different between survey modes within countries using both modes and compared to countries only using CAWI. All well-being variables are strongly impacted by survey mode. On average, CATI tends to produce higher well-being scores. Failure to account for these survey mode differences leads to the reporting of country-level means that are partly an artefact of the way the survey is conducted.

¹ For other examples of differences in the age profile of subjective ill-being and well-being by survey mode see Blanchflower, 2025; Blanchflower and Bryson, 2025a, 2025b, 2025c).

Combined with the presence of other unobservable country-level confounders, the differences between survey modes substantively weakens the internal validity of cross-country comparisons and random effects meta-analysis without controls, a methodology frequently deployed when analyzing GFS.

We examine thirty-three well-being variables, capturing both negative and positive affect, taken from the GFS and twenty-three previous papers using these data.² We first examine a not-flourishing variable from the first sweep of the study (GFS1) first used and defined in Blanchflower and Bryson (2025a). We then replicate the analysis with equivalent data for the second sweep of the survey (GFS2). The four components of not flourishing are also examined separately. They are anxiety, depressed, worry and lack of interest, all of which are coded as 4-step variables. We also look at not being lonely which is a 11-step variable. Then, we examine flourishing, as defined by VanderWeele et al. (2025), from both GFS1 and GFS2, together with its twelve components, all of which are 11-step variables scored on a 0-10 scale. We also examine several other 4-step measures including pain, capable and balance of life and two 5-step variables for religious attendance and community participation. In addition, we examine a number of 1,0 dummy variables including volunteering, donating to charity and health problems.

With the exception of a (not) lonely variable that has not been examined before in the literature, these variables have been previously studied by several other authors using the GFS1. In total there are sixteen separate studies, some of which examine more than one variable. We refer to these sixteen studies henceforth as the GFS secondary analyses. The list of the 33 variables we examine are listed below along with the corresponding studies. **Table 1** provides details of how the various separate variables are defined with the sixteen separate study numbers in parentheses.

- 1) *Not flourishing* (#1) - Blanchflower and Bryson (2025a)
- 2) *Anxious* - Blanchflower and Bryson (2025a)
- 3) *Depressed* - Blanchflower and Bryson (2025a)
- 4) *Worry* - Blanchflower and Bryson (2025a)
- 5) *No Interest* - Blanchflower and Bryson (2025a)
- 6) *Not Lonely*
- 7) *Flourishing* – (#2) VanderWeele et al. (2025) (12 variables).
- 8) *Life satisfaction* - (#3) Lomas, Koga et al (2025).
- 9) *Happiness* - Lomas, Koga et al (2025).
- 10) *Physical health* - (#4) Bradshaw, et al (2025).
- 11) *Mental health* - VanderWeele, et al (2025)
- 12) *Worthwhile* - VanderWeele, et al (2025)
- 13) *Purpose* - (#5) Kim, Noah Padgett, Bradshaw (2025).
- 14) *Promote* - VanderWeele, et al (2025)
- 15) *Give up* - VanderWeele, et al (2025)
- 16) *Content* - VanderWeele, et al (2025)
- 17) *Satisfied with relationships* - VanderWeele, et al. (2025)
- 18) *Expenses* - VanderWeele, et al. (2025)
- 19) *Worry about safety, food and housing* - VanderWeele, et al. (2025)

² <https://globalflourishingstudy.com/research/>

- 20) *Cantril life evaluation* - Lomas, Koga et al (2025)
- 21) *Belonging* - (#6) Counted, et al (2025).
- 22) *Balance of life* - (#7) Lomas, Noah Padgett, Ritchie-Dunham et al. (2025a).
- 23) *Inner peace* - (#8) Lomas, Noah Padgett, Ritchie-Dunham et al, (2025c).
- 24) *A sense of mastery* - (#9) Kim, Noah Padgett, Bradshaw M et al. (2025).
- 25) *Gratitude* - (#10) Okuzono, et al. (2025).
- 26) *In trouble* - (#11) Ritchie-Dunham et al. (2025).
- 27) *Frequency religious attendance* – (#12) Chen et al, (2025).
- 28) *Volunteering* - (#13) Nakamura Woodberry, Gibson et al (2025).
- 29) *Charitable giving* - (#14) Nakamura, Węziak-Białowolska, Woodberry et al (2025).
- 30) *Helping strangers* - Nakamura, Węziak-Białowolska, Woodberry et al (2025).
- 31) *Community participation* - Chen et al, (2025)
- 32) *Health problems* - (#15) Paltzer, Carter, Kwon et al (2025).
- 33) *Bodily pain* - (#16) Macchia, et al (2025).
- 34) *Social support* - Ritchie-Dunham, Yancey, Managi et al (2025).

A second set of papers has been published in journals by broadly the same set of authors using exactly the same dependent variables and the same methodology outlined in Padgett, Bradshaw, Chen et al (2025), but this time the new papers simply add a set of childhood predictors.³ A REMA is also conducted but, once again, these papers do not account for survey mode. We also examine the impact of childhood variables below and find they have only a minor impact on the main variables we consider.⁴ In total we examine a remarkable twenty-three papers that use the GFS1 all of which use the same methodology, many of these papers have the same authors many of whom are members of the Human Flourishing Program at Harvard (<https://hfh.fas.harvard.edu/>).

With the exception of Blanchflower and Bryson (2025a) these studies do not account for survey mode. It turns out that survey mode has a major effect on all of the variables and thus the findings in *every one* of these papers. We highlight results by age, gender, and employment status, but we also report estimates in a **supplementary appendix** for other demographic categories including, educational attainment, being an immigrant, religious attendance, and marital status. We particularly focus on the evidence relating to the young studied in a number of earlier papers including Blanchflower (2025b-g) and Blanchflower, Bryson and Xu (2025).

All of the GFS secondary analyses undertaken by VanderWeele and colleagues take a common methodological approach using random effects meta-analysis (REMA). Firstly, each paper calculates the weighted mean estimates of the variable in question for each country. Next, they treat each of the twenty-two countries as separate studies. Then they estimate a meta-analysis of

³ There are eleven studies that also include childhood predictors - i) (#17) Lomas, Noah Padgett, Ritchie-Dunham et al. (2025) on inner peace; ii) (#18) Lomas, Padgett, Ritchie-Dunham et al. (2025b) on balance of life; iii) (#19) Nakamura, Woodberry, Weziak-Bialowolska et al. (2025) on charitable giving; iv) (#20) Kim, Noah Padgett Bradshaw et al (2025) on purpose and meaning; v) (#16) Macchia, Okafor, Breedlove et al (2025) on physical pain; vi) (#13) Nakamura, Woodberry, Gibson et al (2025) on volunteering; vii) (#11) Ritchie-Dunham, Yancey, Managi et al (2025), on social support; viii) (#21) Weziak-Bialowolska, Cowden, Bialowolski et al (2025) on religious reading; ix) (#22) Kim, VanderWeele & Johnson (2025) on social trust and x) (#23) Paltzer, Kwon, Okafor, et al (2025) on health limitations; xi) Weziak-Bialowolska, Cowden, Paltzer, Bialowolski et al (2025) on finding strength or comfort in your religion or spirituality.

⁴ Adverse child experiences (ACES) are well known to worsen adult well-being (Blanchflower and Bryson, 2024).

the means with a multiplicative random effect term being introduced at the final stage.⁵ Finally, the study authors interpret the country means estimated in this fashion.

We replicate the REMA methodology from each of these studies using STATA 18. We then show that in *every case* the means from all of these studies are highly sensitive to survey mode, that is whether an interviewer is involved (CATI) or if the data is collected through self-reports on the internet (CAWI).⁶ The published studies have overlooked the role of survey mode and thus the mean well-being scores reported are biased. Furthermore, partial correlations between well-being responses and a range of individual characteristics (age, gender, immigrant, labor force, marital status, and religious attendance) differ by survey mode.

The failure to account for survey mode is puzzling since recent published papers emphasize its importance when analyzing responses to questions about individuals' subjective well-being. These include papers using the GFS data (Blanchflower, 2025; Blanchflower and Bryson, 2025a, 2025b, 2025c). For example, Blanchflower and Bryson (2025a) for the Middle East and Asia sample of GFS1 showed that the sign on the age 18-24 variable flipped across eighteen positive and negative affect variables by survey mode.⁷ With CATI the young age 18-24 were, for example, less depressed, and happier than older respondents, but the opposite was the case with CAWI. By contrast, there were no differences on more objective variables such as smoking, drinking, and exercise. One potential reason for these differences is that respondents are more likely to provide positive responses to interviewers than they are online, something that is known as social desirability bias in the psychology and survey methodology literatures (Rickwood and Coleman-Rose, 2023).

We also address a few additional complexities arising from the longitudinal nature of the GFS panel and the survey weights. There are major attrition issues in the second sweep that we examine in a follow-up paper, as the young disproportionately do not show up in wave 2 and there are sharp differences in attrition by survey mode and country (Blanchflower, Bryson and Cui, 2025). The results are also sensitive to weighting as, in a number of countries, the unweighted response rates of the young in particular are very low. The weighting essentially multiplies youth observations upward. In the case of the United States, for example, each respondent aged 18–24 is multiplied by a factor of over nine while those age 65 and over are multiplied by factor of below 0.5. As young people continue to attrit without being replaced and as they age, the GFS will end up with very few observations in the entire 18–24 age group.

The GFS secondary analyses imputed missing values. In none of the studies do the authors explain the impact of the imputation method. It does not appear the imputation has a big effect as there are generally few missing observations. For example, in the case of happiness there were 202,459 observations and for anxiety 202,128 compared to an overall sample of 202,897. We are able to

⁵ See VanderWeele (2017), VanderWeele, Case, Chen et al (2023), Vanderweele McNeely and Koh (2023), Padgett Bradshaw, Chen et al (2025). See also <https://globalflourishingstudy.com/research/>

⁶ Nigeria had 550 observations using CAPI method. A CAPI survey, or Computer-Assisted Personal Interviewing, is a method where interviewers use a digital device like a tablet or smartphone to conduct face-to-face interviews, recording responses in real-time.

⁷ The countries are Hong Kong, India, Indonesia, Israel, Japan, Philippines and Turkey

replicate exactly the authors' results from the REMA without imputation and hence do not impute missing values in our replications.

2. Global Flourishing Study Survey Methodology

The Global Flourishing Study is a panel study with five sweeps planned. Waves 1 and 2, collected principally in 2023 and 2024 respectively, are now available, with three further sweeps planned in 2025 to 2027. We examine data from 22 countries: Argentina; Australia; Brazil; Egypt; Germany; Hong Kong; India; Indonesia; Israel; Japan; Kenya; Mexico; Nigeria; Philippines; Poland; South Africa; Spain; Sweden; Tanzania; Turkey; UK and USA.⁸ Total sample size is 202,897. The second sweep (GFS2) was undertaken across the same countries with a sample size of 124,324 that we also briefly examine. The survey mode used by an individual in the first sweep was repeated in the second. Of the 22 countries seven used only CAWI for data collection (Australia, Germany, Hong Kong, Spain, Sweden, UK, USA) while fifteen use both CAWI and CATI (Argentina, Brazil, Egypt, India, Indonesia, Israel, Japan, Kenya, Mexico, Nigeria, Philippines, Poland, South Africa, Tanzania and Turkey). **Table 2** presents sample size by survey mode in each country.

It would have been possible to evaluate the impact of survey mode on well-being responses if it had been randomly assigned across those eligible for the study, but it was not. Instead, the choice of survey mode for each respondent was determined “*depending on the participant’s access to the internet and willingness to complete online surveys*” (Noah Padgett et al., 2025b).

In addition to differences in survey mode by country, the authors note other limitations affecting the internal validity of cross-country analysis using REMA, including, “*seasonality effects, differences in interpretation of items or response scales, differences due to quality of translation, [...] differences in the process and variables used for constructing respondent level weights, and other possible reasons depending on the specific construct of interest*” (Noah Padgett et al., 2025b). Furthermore, the sampling method differed between countries with some having a probability-based sample (such as phone sampling) and others having a non-probability-based sample collected through web panels or a mix of both sampling methods.

Thus, the design of the GFS varies significantly by country, posing challenges when seeking to make cross-country comparisons in well-being when using REMA. We examine differences in the estimates by survey mode. Further differences between countries due to unobserved confounders may compound the severity of the internal validity issues illustrated by survey mode differences.

3. Why Survey Mode Matters

The method used to elicit responses from those eligible for a survey can have profound implications for the data collected because, as the previous literature indicates, response rates and the nature of responses given can vary across survey modes. For example, people completing interviewer-administered questionnaires are more likely to provide socially desirable responses than in self-administered, internet surveys (Rickwood and Coleman-Rose, 2023; Milton et al.,

⁸ Data is also available on some variables for China, but we exclude it from our analysis as do all the studies we examined.

2017; Aquilino, 1998; Burkill et al, 2016; D'Ancona, 2014).⁹ Rickwood and Coleman-Rose (2023) define *social desirability bias* as “the tendency to under-report socially undesirable attitudes and behaviours and over-report more desirable attributes. This may happen for two reasons: first, for *impression management*, which is the deliberate presentation of self to conform to an audience's normative expectations; and second, due to *self-deception*, which is based on motivation to maintain a positive self-concept that may be unconscious,” (2023, p.2).

Rickwood and Coleman-Rose (2023) looked at two large national community samples of young Australians aged 12–25 years conducted in 2020 and 2022 which used both interviewer-administered and self-report modes of data collection. Results showed participants reported lower psychological distress and higher wellbeing in the interviewer-assisted compared with the self-report mode. Kocjan, Lavtar & Sočan (2023), in a study in Slovenia, reported statistically significant and notable differences in latent means, suggesting that individuals who respond face-to-face systematically report better psychological functioning than individuals who respond over the web.

Who responds to the survey may also be impacted by survey mode. It seems that the young are less likely than they used to be to make or receive phone calls. In an Edison Mail 2022 survey of Gen Z ages 9-24 in Australia, only 23% said that phone calls were their most used communication method versus 33% for those age 33-46 and 62% for ages 76-93.¹⁰ Research by CommBank and More in Australia found only 1 in 10 Gen Z Australians would prefer to talk to their friends and family via a phone call. Overall, 87 per cent admitted to handling an unpleasant task via text message instead of picking up the phone, with nearly half admitting that speaking on the phone makes them feel anxious (49 per cent).¹¹ Thus survey mode may impact the composition of those responding to a survey in a way which may affect the nature of the data collected.

Cernat et al. (2016) look at depression symptoms and Piccitto et al. (2022) look at subjective domain satisfaction and loneliness. Both use a quasi-experimental design and show lower subjective wellbeing in the web sample. In a 2018 study of Croatia, Germany and Portugal Piccitto et al (2022) show that web-interviews provide lower well-being scores among the young, compared to face-to-face interviews and note that “respondents report much lower satisfaction with their job, satisfaction with their partner relationship and subjective health, and higher levels of loneliness”. Cernat et al (2016) examined the US Health and Retirement Surveys 2010-2012 leading them to “caution against mixing interviewer and self-administered modes where possible and encourage study designs that allow for the evaluation of mode effects across a range of topics and indicators.”

Keeter (2015) from the Pew Research Center evaluated mode of interview effects in the United States.¹² Telephone responses were compared to those via the web and “sizeable mode differences

⁹ References from Rickwood and Coleman-Rose (2023).

¹⁰ <https://www.edisonmail.com/blog/study-across-generations-email-remains-a-critical-tool>

¹¹ <https://www.commbank.com.au/articles/newsroom/2023/06/CBA-More-Gen-Z-research.html>

¹² The study used Pew Research Center’s American Trends Panel, a probability-based, nationally representative panel of US adults living in households. Respondents who self-identify as internet users (representing 89% of U.S. adults) participated in the panel via monthly self-administered Web surveys, and those who do not use the internet participate via telephone or mail. All current members of the American Trends Panel were originally recruited from the 2014

were observed on questions measuring satisfaction with family and social life. Among phone survey respondents, 62% said they were “very satisfied” with their family life; among Web respondents, just 44% said this. Asked about their social life, 43% of phone respondents said they were very satisfied, while just 29% of Web respondents said this. These sizeable differences are evident among most social and demographic groups in the survey”.

Respondents in the Pew surveys were also asked about satisfaction with their local community as a place to live. Phone respondents were again more positive, with 37% rating their community as an excellent place to live, compared with 30% of web respondents. But there was no significant difference by mode in the percentage who gave a negative rating to their community (“only fair” or “poor”). Telephone survey respondents were more likely than web respondents to describe their neighborhood as a “very safe” place to walk after dark (55% in the phone survey, 43% in the Web survey).

Another factor potentially contributing to the impact of survey mode is that internet usage may itself be endogenous, affecting outcomes on well-being variables. Several instrumental variable studies using quasi-experimental designs have studied this causal link, each finding that the spread of broadband to a neighborhood is linked to a subsequent increase in ill-being in both subjective self-reports and objective measures, such as hospitalizations for mental health emergencies (Pugno, 2025). In the case of the GFS, survey mode may be linked to internet usage as the authors suggest the choice of survey mode was determined by the respondent’s ability and willingness to use the internet to respond (Padgett et al., 2025b). If these links exist, survey mode may be inadvertently capturing the effect of the internet itself, adding to bias in overall estimates using all modes and weakening the validity of cross-country comparison and the uncontrolled REMA methodology.

While survey mode is a significant consideration in analysis of GFS data, further complications arise as we consider future waves of the GFS, including the second wave which has been released. These complexities arise due to non-random attrition and survey weights, raising additional challenges for secondary analysis of future waves of the GFS.

4. Random Effects Meta Analysis by Survey Mode

All the GFS secondary analyses we examine use REMA, and none took note of the sensitivity of results to survey mode. The **Supplementary Appendix** contains the full results for each of the thirty-three variables. We report four separate results: overall estimates replicating the GFS secondary analyses (column 1), results for CATI (CATI15) and CAWI (CAWI15) for the fifteen countries that use both survey modes (columns 2 and 3 respectively), and results in the seven countries with only CAWI (CAWI7, column 4). Results are reported for seven demographic variables: age (8 categories); gender (3 categories); immigrant (2 categories); marital status (6 categories), education (3 categories); labor force status (7 categories) and religious attendance (5 categories).

For each of the thirty-three variables we examine, we find very different means in age for CATI15, CAWI15 and CAWI7. In many cases, we see different paths in age for each of the three survey

Political Polarization and Typology Survey, a large (n=10,013) national landline and cellphone random digit dial (RDD) survey conducted January 23-March 16, 2014, in English and Spanish.

modes. We plot the means for each of the age groups in **Charts 1–34**. Although there is variance across the well-being metrics, it is apparent how different the means are in the three survey modes and in many cases the shape of the functions varies by age.

Charts 1a and 1b show not flourishing from GFS1 and GFS2, respectively. They look similar. Not-flourishing declines in age but is flatter for the CATI15 than CAWI15 or CAWI7. **Charts 2–5** plot its four components – anxiety, depression, worry and lacking interest. All four are declining in age but the decline is most marked in CAWI7. Levels also look very different with these four aspects of negative affect most apparent for CAWI15. For not flourishing and its components, differences between CATI15 and CAWI15 within the same countries are large and statistically significant for most age groups, and especially large for the young. **Chart 6** shows not lonely rises in age but most steeply for CAWI7.

Charts 7a and 7b are for flourishing from GFS1 and GFS2 respectively and **Charts 8–19** are for the twelve individual components of flourishing. In contrast to not flourishing, flourishing rises in age overall for the two CAWI samples but is roughly flat for CATI15 in GFS1 and is actually declining in age for CATI15 in GFS2.

It is notable how different the three survey mode estimates are the 12 well-being variables used in the flourishing index. **Chart 8** for life satisfaction shows a steep rise in age for CAWI7, a less steep rise for CAWI15 and a shallow U-shape for CATI15. The levels come together for older people, but among the young the levels differences are stark. A roughly similar pattern can be seen in **Chart 9** for happiness. These trends do not correspond with the conclusion by Lomas, Koga, Noah Padgett et al (2025) who conducted a REMA for life satisfaction, and happiness. They argue that "*our results broadly support the well-established view that E-SWB is somewhat "U-shaped," declining into middle-age before rising again.*" (p.18).¹³ They also claim - "*higher E-SWB is associated with older age, though levels are fairly high at 18-24 (6.35/6.78/6.96), falling to their lowest at 40- 49 (6.18/6.70/6.86), then peaking at 80+ (6.83/7.17/7.39).*". That is certainly not the case for CAWI7 for either variable, which rises in age. In both charts means, for example, for age 18-24 differ by survey mode. For life satisfaction the overall mean is 6.78 (6.96), versus 7.09 (7.28) for CATI15, 6.73 (7.07) for CAWI15 and 6.41 (6.51) for CAWI7 with means for happiness, with similar differences in parentheses.

Chart 10 shows respondents' perceptions of their physical health declines in age but the age gradient is much steeper for CATI15 than for CAWI surveys.

In **Charts 11–14** mental health, being worthwhile, life purpose, and promoting good all improve in age but differences in levels are notable, with differences across CAWI and CATI persisting, apart from later in life when CAWI15 respondents post higher scores than CATI15 respondents.

As described in **Table 1**, 'Give up' captures the feeling that an individual always feels able to "give up some happiness now for greater happiness later". **Chart 15** indicates that those least able to give

¹³ The authors say "*there remains uncertainty about how SWB should best be assessed, including regarding its evaluative dimension (E-SWB), with life evaluation (LE) and life satisfaction (LS) the dominant candidates for indexing E-SWB, but with happiness (H) also a possible option (albeit more ambiguously, since in many classifications this is considered a hedonic/affective construct).*" It is unclear to us what that means.

up happiness in the short-term are those in the CAWI7 countries. Differences between CAWI7 and CATI15 diminish with age, whereas ‘give up’ rises with age in CAWI15.

Feeling content and satisfied with relationships slope upwards with age (Charts 16 and 17 respectively), and survey mode differences appear modest within the 15 countries with both CAWI and CATI.

Worrying about expenses (Chart 18) and worry about food, safety and housing (Chart 19) exhibit a shallow U-shape in age, with worry being far more prevalent in the CAWI7 countries across the age range.

Chart 20 shows that Cantril's life evaluation rises with age in the CAWI7. But within the 15 countries with both survey modes, trends in life evaluation diverge, with CAWI scores rising and CATI scores falling.

Charts 21–23 showing belonging, life balance, inner peace rise in age, but at different rates for each survey mode such that, by one's 60s, there is little difference across survey modes. When reporting on belonging by age Counted, Allen, Johnson & VanderWeele (2025: 5) noted that "*older vs. younger participants reported incrementally higher levels of belonging*". But when we break out differences by survey mode, we see that most of this is driven by the age profile of belonging among the CAWI7 countries.

Charts 24 reports mastery. It is roughly flat across age for the CATI15 and CAWI15 but rises for CAWI7. A similar picture emerges for gratitude in Chart 25. Being able to count on others when one is in trouble is more common in the CAWI7 developed countries. From their 30s CAWI15 are more likely to be able to do so than CATI15 respondents.

Charts 27 to 29 capture three behaviors which are often correlated with wellbeing but are not, themselves, wellbeing variables. These are religious attendance (Chart 27), volunteering (Chart 28) and charitable giving (Chart 29). Religious attendance varies little by age irrespective of survey mode, though those in the developed CAWI7 countries are less likely to attend religious services than those in the remaining 15 countries. By contrast, volunteering tends to rise early in life, before tailing off later in life. Charitable giving rises early in life, but remains stable thereafter, with those in CAWI7 countries much more likely to give.

Helping strangers and community participation (Charts 30 and 31 respectively) gradually decline later in life, and do not differ markedly by survey mode.

Health problems and bodily pain both rise in age (Charts 32 and 33 respectively) but much more markedly in the case of CATI15 compared to CAWI15).

Chart 34 for social support shows that it declines gradually with age, with levels higher throughout the age distribution in the CAWI7.

Quite substantial differences in wellbeing persist across survey modes, even within demographic categories. For example, Part a) of Table 4 shows REMA differences by survey mode across males

and females. For both men and women, flourishing is highest, and not-flourishing is lowest in CATI15. Part b) presents REMA estimates for employees and the unemployed. It shows that flourishing is higher for the employed than the unemployed in both CATI15 and CAWI15, but the gap is much larger in the case of CAWI15.

The interpretation of the results when survey mode is accounted for is very different than when it isn't. For example, Counted, Allen, Johnson & VanderWeele (2025) – henceforth CAJV examine belonging. The authors noted (p.5) that "*older vs. younger participants reported incrementally higher levels of belonging (e.g. mean belonging age 80+ = 8.21 [95% CI: 7.86, 8.56] vs. mean age 18–24 = 7.47 [95% CI: 7.08, 7.85])*". We report belonging in [Chart 21](#) overall and by survey mode. It is noticeable how different the means are for the four plots. For example, the means for age 18-24 are 7.47 overall, but 7.71 for CATI15, 7.76 for CAWI15 and 7.01 for CAWI7. CATI15 rises by age, CAWI15 declines in age through age fifty, while CAWI7 is flat through age 40. Neither CAWI15 nor CAWI7 report incrementally higher levels of belonging with age.

In addition, in the [Supplementary Appendix](#) for belonging and every other variable examined here we include results by survey mode, as do the earlier papers, for six other variables – gender, immigrant status, education, employment status, marital status and religious attendance. There are differences among these variables also by survey mode. As an example, with regard to belonging we show below there are marked differences by survey mode for a representative subgroup for each of the seven variables as reported by Counted et al (2025).

	All	CATI15	CAWI15	CAWI7
Age 18-24	7.47	7.71	7.76	7.01
16+ years education	7.68	8.01	7.77	7.39
Immigrant	7.38	7.85	7.30	7.10
Unemployed	7.27	7.77	7.48	6.79
Female	7.77	8.00	7.90	7.58
Divorced	7.60	7.67	7.76	7.52
Never religious attendance	7.36	7.62	7.37	7.13

We find similar differences in the other variables we examine by survey mode reported in that supplement. The issue of bias by survey mode a pervasive phenomenon here.

Different effects by survey mode, both within and between countries, may impact country rankings. One possibility could be, for example, to rank countries using CAWI only, as that was conducted in every country. If we did that here for example using the flourishing variable, there were only small changes in country rankings.¹⁴

Given the impact of survey mode and other unobservable sources of bias, it is understandable that results using the GFS diverge from well-documented determinants of wellbeing, such as the level

¹⁴ Overall ranks are presented here with CAWI data rankings in parentheses. Argentina 10 (16); *Australia 16 (17); Brazil 17 (19); China 11 (13); Egypt 7 (11); *Germany 18 (18); *Hong Kong 13 (14); India 20 (2); Indonesia 1 (1); Israel 2 (7); Japan 23 (23); Kenya 8 (9); Mexico 4 (3); Nigeria 6 (8); Philippines 3 (4); Poland 5 (5); South Africa 15 (10); *Spain 19 (20); *Sweden 14 (15); Tanzania 9 (6); Türkiye 22 (22); *UK 21 (21); *USA 12 (12). Noting that seven countries marked * only have CAWI data.

of development. This divergence is clearly illustrated by the mismatch between country rankings on various development indices and GFS well-being variables, to which we turn now.

5. Country Rankings

In **Table 5**, we rank GFS countries by twenty-one variables of interest. Ranks are based on higher wellbeing so rank 1 is given to a country with the highest happiness level or the lowest unhappiness level. We also construct an overall ranking, GFS summed – presented in column 1 - by simply summing each of the twenty-one ranks and then re-ranking. As was found in Blanchflower and Bryson (2023) country rankings on well-being measures are sensitive to the choice of variable and data set.

When comparing GFS ranks with GNI per capita and the Human Development Index (HDI), the poorer countries tend to have higher well-being. The top ten ranked countries in the GFS are listed below alongside their GNI per capita (2021 PP\$)¹⁵ and HDI rank out of 193 countries:

	Rank	GNI / capita	HDI rank 2023
Indonesia	1	\$13,700	113
Israel	2	\$48,050	27
Mexico	3	\$21,813	81
Nigeria	4	\$5,569	164
Kenya	5	\$5,608	143
Poland	6	\$42,218	35
Argentina	7	\$25,876	47
Philippines	8	\$10,731	117
Hong Kong	9	\$69,436	8
South Africa	10	\$13,694	106

This result seems unlikely given we know from Easterlin (1974) that well-being tends to rise in income up to some threshold. We do not see anything like these results in other surveys. The USA is ranked 11th in the GFS with a GNI of \$73,650 and an HDI rank of 17. Japan ranks lowest. According to the OECD Better Life Index,¹⁶ Japan ranks relatively highly on multiple well-being measures although it does have a relatively high suicide rate.¹⁷ Of particular note is that only one of the top ten ranked countries had data collected solely by CAWI. The high ranking of the countries with low GNI per capita appears in large part to be driven by the fact that they used CATI to draw the data, which tend to report higher levels of well-being. As the impact of survey mode on the REMA estimates and country ranking results has been established, we now turn to reasons why survey mode is such a crucial consideration.

6. Attrition and Weights

The second sweep of the survey, GFS2, has 124,324 observations due to a non-response rate (NRR) of 39%.¹⁸ The attrition patterns contain various worrying features, including differences in attrition by survey mode and increased attrition if respondents reported many scores of ten on

¹⁵ <https://hdr.undp.org/data-center/human-development-index#/indicies/HDI>

¹⁶ <https://www.oecd.org/en/data/tools/well-being-data-monitor/current-well-being.html>

¹⁷ <https://worldpopulationreview.com/country-rankings/suicide-rate-by-country>

¹⁸ We examine attrition rates in a companion paper (Blanchflower, Bryson and Cui, 2025).

the flourishing variable's twelve sub-components. In the case of flourishing, when results are obtained via CAWI, the missing probability rises sharply as the number of answers with a score of ten increases. We calculated the unweighted proportion of respondents who reported a ten to each of the twelve components of flourishing and summed them and we report the results below. The mean was 2.74. Overall, 37% of the sample had zero tens, 13% had one and 11% had three while 6% had 10-12.¹⁹ Overall 38% did not respond to the second survey. Of those who had zero tens 35% did not respond to the second sweep and was higher for CATI (44%) than CAWI (35%). The rise in the missing proportion is most apparent using CAWI. Fifty-eight percent of the CAWI sample with twelve tens attrited. In the CAWI sample the attrition rate rises with the number of tens but is broadly flat for CATI.

# Tens	0	5	8	9	10	11	12	All
All	35	40	42	43	47	46	49	38
CATI	46	44	43	44	46	46	45	44
CAWI	34	37	39	46	49	49	58	35

On the other hand, in the case of not flourishing – which ranges from 4 to 16 - a higher score in GFS1 predicts a higher probability of attrition. Thus, a score of 4=30%; 8=41%; 12=47%; 16=50% and overall=39%. It seems unusual that attrition rates would be higher for those with higher wellbeing.

Users of the GFS data tend to impose weights supplied with the data. But these do not help address the shrinkage in the underlying unweighted sample sizes for some subgroups like the young. This is going to be a major issue going forward as the intention is that there will be three further surveys. In GFS2, by construction there are no young people who are age 18 as the sample has not been replenished as is usual in panel analyses such as the US Health and Retirement Survey. The eighteen-year-olds in sweep one are age nineteen in sweep 2.

To illustrate the problem let us take the case of the United States. All responses in the United States were self-reports obtained via the internet. In GFS1 (n=38,312), the US had 296 young people aged 18-24 or 0.77% of the total. Once weights were imposed each observation is multiplied by nine to 7.00%. This is well below the 11.7% of the 18+ population that is reported by the Census Bureau in its International Database.²⁰ The concern also is that young people tend to have high probabilities of attrition in subsequent panel rounds. This is the case here. It turns out that in sweep 2 for the US there were only 143 age 19-24 and a total of 32,003 observations.²¹ If this non-response rate of 52% repeats subsequently then sample size will be only 20 young people in the US sample at the fifth sweep with the youngest respondent being 23 years old.

	Sweep 1		2025
	Unweighted	Weighted	Census
<25	0.77	7.00	11.7%

¹⁹ The overall distribution of tens was as follows (n= 202,351) - 0=37.21%; 1=13.03%; 2=11.30%; 3=7.60%; 4=6.09%; 5=5.00%; 6=4.36%; 7=3.66%; 8=3.15%; 9=2.72%; 10=2.65%; 11=1.52% and 12=1.71%

²⁰ <https://www.census.gov/programs-surveys/international-programs/about/idb.html>

²¹ Age 18 n=8 (0), 19 n=13 (6), 20 n=21 (10), 21 n=33 (9), 22 n=49 (28), 23 n=64 (33), 24 n=108 (57). Hence 18-24 n=296 (143).

25-29	2.85	9.24	8.5%
30-39	11.21	19.01	17.3%
40-49	11.09	14.74	15.9%
50-59	17.66	17.60	15.0%
60-69	26.75	17.83	15.2%
70-79	23.09	10.58	10.8%
80+	6.57	3.98	5.5%

Altogether, we have discussed why survey mode differences, combined with other possible unobservable country-level confounders may undermine the internal validity of the cross-country comparison and uncontrolled REMA models used by the GFS secondary analyses. We have also briefly discussed additional complexities arising from not replacing respondents who do not respond in subsequent waves of the GFS. Using regression analysis with country-level fixed effects to control for the country level biases is more appropriate for cross-sectional analysis of the first wave of the GFS.

7. Regression Analysis

When examining cross-country evidence on well-being REMA treats each country observation as a separate study. Whilst common in some fields of academic endeavor, it is rarely used in fields such as economics to examine cross-country differences in outcomes such as GDP, unemployment, wages and well-being (El Jahel, MacCulloch and Shafiee, 2024). Instead, the standard approach is to run regression analyses with and without controls, including country/state fixed effects to adjust for permanent and unchanging characteristics.²² When there are multiple years of data, model specifications often incorporate year fixed effects alongside country fixed effects, and may also incorporate time-varying co-variables such as macro-conditions captured by the unemployment rate or the inflation rate.²³

In the eleven papers listed earlier, childhood variables were examined. There is a large pre-existing literature on childhood experiences as noted by the Centers for Disease Control and Prevention (CDC) (<https://www.cdc.gov/aces/about/index.html>) and there is nothing much new here.²⁴ The method used in these papers though was somewhat unusual. For example, Kim, Bradshaw, Padgett (2025) note that *"within each country, we fit a weighted linear regression model with complex survey-adjusted standard errors to assess the associations between composite purpose and meaning and all 11 aforementioned childhood correlate variables, simultaneously. Following these individual country analyses, we used a random effects meta-analytic approach to aggregate the regression coefficients across all 22 countries"*.

A simpler alternative in our view is to estimate regressions with exogenous controls for age, gender and country. In our baseline specification we do not include potentially endogenous controls such as marital status and religion. In a second model specification we incorporate the childhood variables listed in the notes to Table 6, and employment status variables. The latter may suffer from endogeneity.

²² The inclusion of country fixed effects wiped out the entire aggregate time series literature in labor economics (Card, 1995).

²³ For a recent example incorporating average temperature by area and time see Tang, Powdthavee and Macchia, 2025.

²⁴ Blanchflower and Bryson (2024) showed that adverse childhood experiences (ACEs) have long lasting effects in adulthood, including on wellbeing.

There are two panels in [Table 6](#). Panel (a) estimates the correlates of not-flourishing – the composite variable summing responses from 4 variables described above which runs from 4 through 16 – while panel (b) estimates flourishing, a composite variable consisting of 12 variables, also described above, with a scale running from 0 to 10 - the twelve components each ranging from zero to ten are totaled and then the sum is divided by 12.

The estimates in [Table 6](#) are confined to those 15 countries which surveyed individuals via CATI or CAWI. We report separate regressions for CAWI and CATI respondents to examine differences in the correlates of not-flourishing and flourishing across survey modes.

If we turn first to the not-flourishing estimates in [Table 6a](#) it is apparent from columns 1 and 2 that, among CATI respondents, not-flourishing rises with age until people reach their mid-50s, before declining again. This hump-shape has been observed previously in the literature. In GFS the addition of controls in column 2 makes no real difference to this pattern. Columns 3 and 4 run the same regressions, but this time for people who responded via CAWI. The age pattern in not-flourishing is now totally different: it declines monotonically with age until respondents reach their 80s. Again, the addition of childhood and labor market status controls makes no difference. The contrast in the age pattern of not-flourishing by survey mode is remarkable but confirms findings in an earlier paper (Blanchflower and Bryson, 2025a).

[Table 6b](#) runs identical analyses, but this time for flourishing. In the CATI models flourishing declines with age until respondents hit their mid-50s or early 60s whereas, for CAWI respondents, flourishing rises with age once people hit their 30s. So, once again, results look utterly different by survey mode. Furthermore, in contrast to not-flourishing where women had positive significant coefficients regardless of survey mode, in the case of flourishing women have slightly higher flourishing scores than men when interviewed via CATI but lower flourishing scores than men when interviewed via CAWI.

[Table 7](#) repeats the exercise but this time for the seven CAWI only countries. As in the case of the countries using both survey modes, where CAWI is used to survey people we see not-flourishing falling with age while flourishing rises with age.

In the [Supplementary Appendix 1](#), we report the full regression results for flourishing, not flourishing, and repeat the same regressions for a further thirteen variables.²⁵ Survey mode persists as an important consideration for the thirteen other variables. For example, in the regression of the comfort and capable variable in the 15 countries with both modes shown by [Appendix Table 1e and 1g](#), the coefficients on age groups double in magnitude and statistical significance in CAWI compared to CATI. For comfort, the coefficients on gender are stronger in CAWI, but not for capable. A similar pattern emerges for childhood predictors such as family income, where the effect is stronger in CAWI. There are also opposing phenomena, such as for parents married, where the effect is double in strength in CATI. These differences in magnitude and statistical power did not emerge because of differences in N, as both CATI and CAWI have around 60,000 observations in these models. Differences between CATI and CAWI are particularly acute for life

²⁵ The variables are comfort, capable, inner peace, balance of life, donated, sense of purpose, pain, volunteer, social trust, health problems, religious reading, trouble and feelings of being close to.

purpose, where the sign changes between the modes and there is a difference of more than 1.3 on an 11-point scale, as shown in the Supplementary Appendix and below: where coefficients and t-statistics are presented.

	CATI15	CATI15	CAWI15	CAWI15
25-29	-.069 (1.93)	-.073 (1.99)	.128 (3.28)	.110 (2.84)
30-39	-.073 (2.34)	-.059 (1.86)	.248 (7.58)	.229 (7.02)
40-49	-.149 (4.30)	-.101 (2.85)	.337 (9.94)	.314 (9.24)
50-59	-.229 (5.75)	-.188 (4.58)	.523 (14.73)	.487 (13.62)
60-69	-.287 (6.06)	-.250 (5.13)	1.039 (27.36)	.911 (23.87)
Childhood	No	Yes	No	Yes

Interestingly, for inner peace and life balance in [Appendix Table 1i and 1k](#), the effect of gender is statistically significant in the CATI models and CAWI model excluding childhood predictors, but the effect disappears in the CAWI model with childhood predictors. Gender is a significant factor for donations for CATI respondents, but not for either of the CAWI models in [Appendix Table 1m](#). Given gender's significant role in many other studies, its lack of significance in some CAWI estimations is an important and surprising nuance that would have been missed had survey mode not been considered.

In further analyses we ran similar estimates for 33 wellbeing outcomes. We collapsed the age variable into a single dummy variable scoring '1' when aged 18-24, zero otherwise. [Table 8](#) shows the signs and significance of age 18-24 variable by survey mode in these regressions, where the controls are identical to those used in column 1 in [Table 6](#). If one focuses on columns 1 and 2, we see that the youth dummy variable flips sign in all but 6 equations, the exceptions being Cantril's life evaluation ladder, trouble, community, donate, health problems, and pain.

8. Conclusions

A major concern with the twenty-three previous papers using the GFS1 data we examined here is that they do not take account of survey mode. This brings many, if not all, of their findings into question. We also question their methodology which treats each of the twenty-two countries as if each is a separate study. We follow a much more well-established procedure in economics – which is used to dealing with cross-country data that estimates a set of pooled country regressions that control for country fixed effects. We know of no other studies that examine cross-country data from a single cross-section as a meta-analysis.²⁶

²⁶ There has been much previous work, especially in sociology and economics, examining data across countries, often with time components, including for example, the Latinobarometers, Asia Barometers, Afrobarometers, the International Social Survey Programme, the European Quality of Life Survey, the World Values surveys, the European Social Survey (ESS), the Gallup World Poll, the Eurobarometer surveys all used in Blanchflower (2021). These have usually been examined via a regression framework with controls for country and year effects. For example, Venetoklis (2019) examined life satisfaction and happiness data for 16 countries from seven rounds of the ESS from 2002 to 2014 and included country and sweep effects. Carlsen and Leknes (2025), examined the Eurobarometers. Macchia and Oswald (2025, 2021) and Macchia (2022), one of the co-authors of the papers we examine here examined well-being data from the Gallup World Poll and estimated well-being regressions that include country and year effects as we do here. Analogously, when examining US data, on macro-variables such as wages or unemployment or well-being, for example, it is appropriate to control for state and year effects (see Blanchflower, Bryson and Sperling, 2024 for wages). A recent paper modelling mortality in the US estimates county*year mortality change equations with state fixed effects - see Foote, Meara, Skinner and Stewart (2025).

We examined a not-flourishing variable with its four 4-step components scored from 1 to 4 and a flourishing variable with twelve 11-step components scored 0 to 10. We also analyzed a further seventeen variables: 1) Cantril's life evaluation, 2) grateful, 3) bodily pain, 4) belonging, 5) trouble, 6) balance of life, 7) inner peace, 8) being capable, 9) being grateful, 10) frequency of religious attendance, 11) volunteering, 12) charitable giving, 13) helping strangers, 14) health problems, 15) community participation 16) social support and 17) being not lonely.

We replicated the work using these variables and the GFS data file in twenty-three earlier papers that use these data. We used random effects meta-analysis (REMA) models as used in the GFS secondary analyses to identify how these 34 variables, differed by survey mode across age, gender, labor force status, immigration status, educational attainment, marital status, and religious attendance. We focus on the age profile of wellbeing given the attention this has received recently due to the declining mental health of the young (Blanchflower, Bryson and Xu, 2025, and Blanchflower and Bryson, 2025a-g). Well-being tended to be higher in the CATI surveys in the fifteen (mostly developing) countries that had both CATI and CAWI. The association between age and wellbeing varied by survey mode. Among CAWI respondents' well-being rises with age, whereas among CATI respondents it tends to rise. The differences between survey modes are often statistically significant, especially for the young. In addition to differences by age group, we also show the impact of survey mode exists for many of the other demographic categories. The presence of bias from survey mode and other unobservable factors explains conflicting results with the remaining literature, as shown by country rankings, and weakens the validity of cross-country comparison and uncontrolled REMA for GFS panel analysis.

We argued our preferred methodology is to estimate a series of regressions that contained country fixed and unchanging permanent effects, following a large literature in economics. We found for example, that age and gender effects were little impacted by the presence of childhood variables as had been used in eleven other studies using the GFS. Even in the regression equations the relation between age and not flourishing was upward sloping in age with CATI and downward sloping with CAWI. The reverse was the case with flourishing. In the regression framework we estimated separate equations by survey mode.

We detail reasons why survey mode matters, including social desirability bias and endogeneity of internet usage, and address additional complexities arising from attrition in future waves of the GFS and survey weights. Given our findings, our preferred method is estimating flourishing and not-flourishing equations using regression analysis with country fixed effects, where we found similar differences by survey mode. If a REMA analysis approach is to be used, we show it is crucial to show how sensitive results are to the method used to collect the data, either by interviewer or via the internet. Failure to do so inevitably biases the results.

References

- Aquilino WS (1998), 'Effects of interview mode on measuring depression in younger adults', *Journal of Official Statistics*, 14 (1): 15–29.
- Blanchflower DG (2025), 'Declining youth well-being in 167 UN Countries. does survey mode, or question matter?', NBER Working Paper #33415, January. <https://doi.org/10.3386/w33415>
- Blanchflower DG (2021). 'Is happiness U-shaped everywhere? Age and subjective well-being in 145 countries', *Journal of Population Economics*, <https://doi.org/10.1007/s00148-020-00797-z>
- Blanchflower DG and Bryson AJ (2025a), 'Not flourishing', NBER Working Paper #34321, October. <https://doi.org/10.3386/w34321> #1
- Blanchflower DG and Bryson AJ (2025b), 'The mental health of the young in Asia and the Middle East: The importance of self-reports', *Quality and Quantity*. <https://doi.org/10.1007/s11135-025-02264-x>
- Blanchflower DG and Bryson AJ (2025c), 'Life satisfaction in Western Europe and the gradual vanishing of the U-shape in age', *Academia Mental Health and Well-Being*, 2(3). <https://doi.org/10.20935/MHealthWellB7877>
- Blanchflower DG and Bryson AJ (2025d), 'The mental health of the young in Eastern Europe and Central Asia', *Review of Economic Development*, forthcoming.
- Blanchflower DG and Bryson AJ (2025e), 'The mental health of the young in Latin America', *Social Indicators Research*. <https://doi.org/10.1007/s11205-025-03669-9>
- Blanchflower DG and Bryson AJ (2025f), 'The mental health of the young in Africa', *Journal of African Economics*, forthcoming.
- Blanchflower DG and Bryson AJ (2025g), 'Rising young worker despair in the United States', *SSM – Population Health*. <https://doi.org/10.1016/j.ssmph.2025.101881>
- Blanchflower DG and Bryson AJ (2024), 'The adult consequences on wellbeing of abuse and neglect in childhood', *International Journal of Wellbeing*, 2024. <https://doi.org/10.5502/ijw.v14i3.3513>
- Blanchflower DG and Bryson AJ (2023), 'Wellbeing rankings', *Social Indicators Research*. <https://doi.org/10.1007/s11205-023-03262-y>
- Blanchflower DG, Bryson AJ and Cui A (2025), 'Attrition bias in the Global Flourishing Study Panel', Working paper.
- Blanchflower DG, Bryson AJ and Spurling J (2024), 'The wage curve after the Great Recession', *Economica*, 91(362): 653-668. <https://doi.org/10.1111/ecca.12515>

Blanchflower DG, Bryson AJ and Xu X (2025), 'The declining mental health of the young and the global disappearance of the hump shape in age in unhappiness', *Plos One*.
<https://doi.org/10.1371/journal.pone.0327858>

Bradshaw M, Kent BV, Levin J, Wortham JS, Pertel NL, VanderWeele TJ, Johnson BR (2025), 'Demographic variation in self-rated physical health across 22 countries: findings from the Global Flourishing Study', *BMC Global Public Health*, Apr 30;3(1): 38. #4
<https://doi.org/10.1186/s44263-025-00141-1>

Burkill S, Copas A, Couper MP, Clifton, Prah P, Datta J, Conrad F, Wellings K, Johnson AM, Erens B (2016), 'Using the web to collect data on sensitive behaviours: a study looking at mode effects on the British national survey of sexual attitudes and lifestyles', *PLoS One* 11(2).
<https://doi.org/10.1371/journal.pone.0147983>

Card D (1995), 'The wage curve: A review', *Journal of Economic Literature*, 33(2): 785-799.
<https://www.jstor.org/stable/2729028>

Carlsen F and Leknes S (2025), 'Who are satisfied with life in cities? Evidence for 25 European countries', *Urban Studies*, 62(11): 2348-2368. <https://doi.org/10.1177/00420980241310375>

Case B, Counted V and Ritchie-Durham et al (2025), 'Beyond a single story: The heterogeneity of human flourishing in 22 countries', *International Journal of Wellbeing*, 13(4).
<https://doi.org/10.5502/ijw.v13i4.3555>

Cernat A, Couper MP, & Ofstedal MB (2016), 'Estimation of mode effects in the Health and Retirement Study using measurement models', *Journal of Survey Statistics and Methodology*, 4(4), 501-524. <https://doi.org/10.1093/jssam/smw021>

Chen Y, Case BW, Long KNG et al (2025), 'Community participation in secular and religious contexts across sociodemographic groups in 22 countries', *Scientific Reports*, 15. #12
<https://doi.org/10.1038/s41598-025-22806-5>

Counted V, Allen KA, Johnson BR & VanderWeele TJ (2025), 'Variations in belonging across 22 countries in the Global Flourishing Study', *The Journal of Positive Psychology*. #6
<https://doi.org/10.1080/17439760.2025.2569083>

D'Ancona MAC (2014), 'Measuring xenophobia: Social desirability and survey mode effects', *Migration Studies*, 2(2): 255–280. <https://doi.org/10.1093/migration/mnt014>

El-Jahel L, MacCulloch R and Shafiee H (2023), 'How does monetary policy affect welfare? some new estimates using data on life evaluation and emotional well-being', *Journal of Money, Credit, and Banking*, 55(8): 2001-2025. <https://doi.org/10.1111/jmcb.13000>

Foote CL, Meara E, Skinner JS and Stewart LR (2025), 'Explaining the widening divides in US midlife mortality: is there a smoking gun?', NBER Working Paper #34553 <http://www.nber.org/papers/w34553>

Keeter S (2015), 'From telephone to the web: the challenge of mode of interview effects in public opinion polls', Pew Research Center.

Kim ES, Noah Padgett, R, Bradshaw M *et al.* (2025), 'Mapping demographic variations in sense of mastery across the world a cross-national analysis of 22 countries in the global flourishing study', *Scientific Reports*, 15, 29904. #9 <https://doi.org/10.1038/s41598-025-15304-1>

Kim ES, Bradshaw M, Noah Padgett R *et al* (2025), 'Identifying childhood correlates of adult purpose and meaning across 22 countries (Global Flourishing Study)', *npj Mental Health Res* 4, 14 (2025). #20 <https://doi.org/10.1038/s44184-025-00127-9>

Kim ES, Noah Padgett R, Bradshaw M *et al* (2025), 'Mapping demographic variations in purpose and meaning across the world: a cross-national analysis of 22 Countries in the Global Flourishing Study', *Scientific Reports*, 15, 29904. #5 <https://doi.org/10.1038/s41598-025-15304-1>

Kim YI, VanderWeele TJ. and Johnson B (2025), 'Childhood predictors of perceptions of social trust across 22 countries in the Global Flourishing Study', *Scientific Reports* 15. #22 <https://doi.org/10.1038/s41598-024-78201-z>

Lomas T, Noah Padgett R, Ritchie-Dunham JL *et al.* (2025a), 'Demographic variation in balance in life across 22 countries: a cross-national analysis of the Global Flourishing Study', *Applied Research in Quality Life* 20, 1011–1036. #7 <https://doi.org/10.1007/s11482-024-10407-9>

Lomas T, Noah Padgett R, Ritchie-Dunham JL *et al.* (2025b), 'Childhood predictors of balance in life: a cross-national analysis of the Global Flourishing Study', *Scientific Reports*, 15. #18 <https://doi.org/10.1038/s41598-025-89853-w>

Lomas T, Noah Padgett R, Ritchie-Dunham JL *et al.* (2025c), 'Demographic Variation in Inner Peace Across 22 Countries: A Cross-National Analysis of the Global Flourishing Study', *Journal of Happiness Studies*, 26, 66. #8 <https://doi.org/10.1007/s10902-024-00822-y>

Lomas T, Koga HK, Noah Padgett *et al* (2025), 'Life evaluation, life satisfaction, and happiness: assessing inter-relations and 15 childhood and demographic factors across 22 Countries in the Global Flourishing Study'. Working paper. #3

Lomas T, Noah Padgett R, Ritchie-Dunham JL *et al.* (2025), 'An exploratory cross-national analysis of the childhood predictors of inner peace in the Global Flourishing Study', *Scientific Reports*, 15. #17 <https://doi.org/10.1038/s41598-024-83353-z>

Macchia L (2022), 'Pain trends and pain growth disparities, 2009-2021', *Economics and Human Biology*, 47: 101200. <https://doi.org/10.1016/j.ehb.2022.101200>

Macchia L, Okafor CN, Breedlove T et al. (2025), 'A cross-national analysis of childhood predictors of physical pain', *Community Medicine*, 5, 337. #16
<https://doi.org/10.1038/s43856-025-00997-2>

Macchia L and Oswald AJ (2025), 'National worry and the psychological value of social spending', *The Review of Income and Wealth*, 71(4), November. <https://doi.org/10.1111/roiw.70039>

Macchia L and Oswald AJ (2021), 'Physical pain, gender, and the state of the economy in 146 nations', *Social Science & Medicine*, 287. <https://doi.org/10.1016/j.socscimed.2021.114332>

Milton AC, Ellis LA, Davenport TA, Burns JM and Hickie IB (2017), 'Comparison of self-reported telephone interviewing and web-based survey responses: findings from the second Australian young and well national survey', *JMIR Mental Health*, 4 (3). <https://doi.org/10.2196/mental.8222>

Nakamura JS, Węziak-Białowolska D, Woodberry RD et al. (2025), 'Demographic variation in charitable giving and helping across 22 countries in the Global Flourishing Study', *Scientific Reports*, 15. #14 <https://doi.org/10.1038/s41598-025-96009-3>

Nakamura JS, Woodberry RD, Gibson CB et al. (2025), 'Childhood predictors of volunteering across 22 countries in the Global Flourishing Study', *Scientific Reports*, 15. #13 <https://doi.org/10.1038/s41598-024-81639-w>

Nakamura JS, Woodberry RD, Weziak-Bialowolska D et al. (2025), 'Childhood predictors of charitable giving and helping across 22 countries in the Global Flourishing Study', *Scientific Reports*, 15. #19 <https://doi.org/10.1038/s41598-024-77950-1>

Okuzono SS, Cowden RG, Padgett N et al. (2025), 'sociodemographic variation in gratitude using a cross-national analysis with 22 countries', *International Journal of Applied Positive Psychology*, 10, 60. <https://doi.org/10.1007/s41042-025-00254-w>

Padgett RN, Bradshaw M, Chen Y et al. (2025a), 'Analytic methodology for demographic variation analyses for wave 1 of the Global Flourishing Study', *BMC Global Public Health*, 3, 28. #10 <https://doi.org/10.1186/s44263-025-00140-2>

Padgett RN, Cowden RG, Chattopadhyay M et al. (2025b), 'Survey sampling design in wave 1 of the Global Flourishing Study,' *European Journal of Epidemiology*, 40(4), pp. 391–406. <https://doi.org/10.1007/s10654-024-01167-9>.

Paltzer J, Carter EW, Kwon E et al. (2025), 'A cross-sectional study of demographic variation in health problem-related limitations in life across 22 countries: a cross-national analysis. *BMC Global Public Health* 3, 78. #15 <https://doi.org/10.1186/s44263-025-00190-6>

Paltzer J, Kwon E, Okafor. et al. (2025), 'Childhood predictors of health limitations in life across 22 countries: a cross-national and cross-sectional analysis', *BMC Global Public Health* 3, 70. #23 <https://doi.org/10.1186/s44263-025-00188-0>

Piccitto G, Liefbroer AC and Emery T (2022), 'Does the survey mode affect the association between subjective well-being and its determinants? An experimental comparison between face-to-face and web mode', *Journal of Happiness Studies*, 23(7), 3441-3461.

<https://doi.org/10.1007/s10902-022-00553-y>

Pugno, M. (2025), 'Does social media harm young people's well-being? A suggestion from economic research', *Academia Mental Health and Well-Being*, 2(1).

<https://doi.org/10.20935/MHealthWellB7581>

Rickwood D and Coleman-Rose CL (2023), 'The effect of survey administration mode on youth mental health measures: Social desirability bias and sensitive questions', *Heliyon* 9(9), e2013.

<https://doi.org/10.1016/j.heliyon.2023.e20131>

Ritchie-Dunham JL, Yancey G, Managi S et al. (2025), 'Childhood predictors of social support and intimate friends in a Cross-National analysis of the Global Flourishing Study', *Scientific Reports*, 15. #11 <https://doi.org/10.1038/s41598-025-09953-5>

Tang CK, Powdthavee N and Macchia L (2025), 'Rising temperatures and physical pain: Evidence from over 2 million U.S. residents', *Ecological Economics*, 241.

<https://doi.org/10.1016/j.ecolecon.2025.108871>

VanderWeele TJ (2017), 'On the promotion of human flourishing', *PNAS*, 114 (31): 8148-8156.

<https://doi.org/10.1073/pnas.1702996114>

VanderWeele TJ, Case BW, Chen Y, Cowden RG, Johnson B, Lee MT, Lomas T and Long KG (2023), 'Flourishing in critical dialogue', *SSM - Mental Health*, 3.

<https://doi.org/10.1016/j.ssmmh.2022.100172>

VanderWeele TJ, Johnson BR, Bialowolski PT et al (2025), 'The Global Flourishing Study: Study profile and initial results on flourishing', *Nature Mental Health*, 3: 636–653. #2

<https://doi.org/10.1038/s44220-025-00423-5>

VanderWeele TJ, McNeely E and Koh HK (2019), 'Reimagining health - flourishing', *JAMA*, May 7, 321(17). <https://doi.org/10.1001/jama.2019.3035>

Venetoklis T. (2019), 'Do interactions cancel associations of subjective well-being with individual-level socioeconomic characteristics? An exploratory analysis using the European Social Survey', *Quality & Quantity*, 53(6): 3033–3061. <https://doi.org/10.1007/s11135-019-00919-0>

Weziak-Bialowolska D, Cowden RG, Bialowolski P et al (2025), 'Childhood predictors of religious reading: a cross-national analysis in the Global Flourishing Study', *Scientific Reports*, 15. #21 <https://doi.org/10.1038/s41598-025-10142-7>

Table 1. GFS Wellbeing Metrics and the Underlying Survey Questions

1) *Not flourishing*

Q1. *Anxious* - feeling nervous, anxious or on edge – 1=not at all, 2=several days, 3=more than half, 4=nearly every day.

Q2. *Depressed* - Feeling down, depressed or hopeless– 1=not at all, 2=several days, 3=more than half, 4=nearly every day.

Q3. *Worry* - Not being able to stop or control worrying– 1=not at all, 2=several days, 3=more than half, 4=nearly every day.

Q4. *No Interest* - little interest or pleasure in doing things– 1=not at all, 2=several days, 3=more than half, 4=nearly every day.

Q5. *Not flourishing* is (Anxious+Depressed+Worry+No interest)/4.

2) *Flourishing*

Q6. *Life* - Overall, how satisfied are you with life as a whole these days? 0=not at all satisfied..10=completely satisfied.

Q7. *Happy* - In general, how happy or unhappy do you usually feel? 0=extremely unhappy ..10=extremely happy.

Q8. *Physical* - In general, how would you rate your physical health? 0=poor...10=excellent.

Q9. *Mental* - How would you rate your overall mental health? 0=poor...10=excellent.

Q10. *Worthwhile* - Overall, to what extent do you feel the things you do in your life are worthwhile? 0=not at all...10=completely.

Q11. *Purpose* - I understand my purpose in life. 0=strongly disagree..10=strongly agree.

Q12. *Promote* - I always act to promote good in all circumstances, even in difficult and challenging situations. 0=not true at all...10=completely true of you.

Q13. *Give up* - I am always able to give up some happiness now for greater happiness later. 0=not true at all...10=completely true of you.

Q14. *Content* - I am content with my friendships and relationships. 0=strongly disagree..10=strongly agree.

Q15. *Satrel* - My relationships are as satisfying as I would want them to be. 0=strongly disagree..10=strongly agree.

Q16. *Expenses* - How often do you worry about being able to meet normal monthly living expenses? 0=worry all the time..10=do not ever worry.

Q17. *Foodhse* - How often do you worry about safety, food or housing? 0=worry all the time..10=do not ever worry.

Q18. *Flourishing* is the sum of (Life satisfaction +Happiness+Physical+ Mental

3) *Not lonely*

Q19. How Often do You Feel Lonely – 1=always....10=never

4) *Life evaluation*

Q20. Life evaluation - today 0=worst possible 10=best possible.

5) *Belonging*.

Q21. ‘How would you describe your sense of belonging in your country?’. The response was rated from 0 (very weak) to 10 (very strong).

6) *Social support*.

Q22. “Is there any one special person you know that you feel very close to? For example, someone you can confide in and share your feelings with. (If you have more than one person who you feel very close to select ‘Yes.’)” [Y/N]

7) *Balance of life*

Q23. “In general, how often are the various aspects of your life in balance?” Response options are: always, often, rarely, never.

8) *Purpose and meaning*

Q24. Purpose and meaning in life. Purpose was assessed by asking “I understand my purpose in life.” Response options ranged from: 0 = Strongly Disagree to 10 = Strongly Agree. Meaning was assessed by asking “Overall, to what extent do you feel the things you do in your life are worthwhile?” Response options ranged from 0 = Not at All Worthwhile to 10 = Completely Worthwhile.

9) *Inner peace*

Q25. In general, how often do you feel you are at peace with your thoughts and feelings? - always, often, rarely, never. In our analyses, they dichotomized IP as always/often [1] vs. rarely/never [0].

10) *Health problems* Paltzer, Carter, Kwon et al. 2025.

Q26. Health problems prevent you from doing things people your age normally Yes/No

9) *A sense of mastery*

Q27. “How often do you feel very capable in most things you do in life?” Response options included: Always, often, rarely, never. In our main analyses, mastery was dichotomized as 1 (always/often) versus 0 (rarely/never)—as described in our pre-registered analyses. We initially chose this cut-off because it reflects a conceptual distinction where individuals who consistently feel capable (“always/often”) are separated from those who have a pronounced deficit (“rarely/never”).

10) *Bodily pain*

Q28. ‘How much bodily pain have you had during the past 4 weeks?’ Respondents could answer a lot, some, not very much, or none at all. In the main analyses, this variable was dichotomised as a lot/some (1) vs. not very much/none at all (0)

11) *Gratitude*

Q29. “If I had to list everything that I felt grateful for, it would be a very long list.” The decision to select this item was informed by prior validation work, input from subject experts, and considerations of cross-cultural applicability. Cognitive interviews and pilot testing were used to ensure clarity and cultural relevance across the countries included in the GFS Response options were anchored at 0 (Strongly disagree) and 10 (Strongly agree). We used this indicator as a continuous variable in the analysis.

12) *Physical health*

Q30. “In general, how would you rate your PHYSICAL health?” Respondents were instructed to rate their health on a 0–10 scale, where 0 = poor and 10 = excellent. This measure was treated as continuous.

13) *In trouble*

Q31. “If you were in trouble, how often could you count on people in your life, like relatives or friends, to help you whenever you need them?” [0 = Never, 10 = Always], assessed as a continuous outcome

14) *Charitable giving*

Q32. “In the past month, have you donated money to a charity binary yes/no for both items.

15) *Helping strangers*

Q33. "Helping strangers was assessed with a single item, which asked, "In the past month, have you helped a stranger or someone you didn't know who needed help?" Response options were binary yes/no for both items.

16) *Community participation*

Q34. "How often do you participate in groups that are not religious, such as book clubs, sports, or political organizations?". Response options include "never", "a few times a year", "one to three times a month", "once a week", and "more than once a week". Because prior studies suggest that at least once/week of community participation is often associated with improved health and well-being outcomes^{17,42,58}, we dichotomized the responses as at least once/week vs. less than once/week or never.

17) *Frequency of religious attendance*

Q35. "How often do you attend religious services?". Response options include "never", "a few times a year", "one to three times a month", "once a week", and "more than once a week". Similar to secular community participation (see above), we dichotomized religious service attendance as at least once/week vs. less than once/week or never.

18) *Volunteering*

Q36. Volunteering was assessed with a single item, which asked, "In the past month, have you volunteered your time to an organization?" Response options were binary (yes/no).

Table 2. Responses by country and survey mode unweighted.

	CAPI	CATI	CAWI	All
Argentina	0	1,451	5,273	6,724
Australia	0	0	3,844	3,844
Brazil	0	1,265	11,938	13,203
Egypt	0	4,508	221	4,729
Germany	0	0	9,506	9,506
Hong Kong	0	0	3,012	3,012
India	0	12,549	216	12,765
Indonesia	0	3,620	3,372	6,992
Israel	0	2,743	926	3,669
Japan	0	542	20,001	20,543
Kenya	0	9,917	1,472	11,389
Mexico	0	1,559	4,217	5,776
Nigeria	550	5,609	668	6,827
Philippines	0	4,174	1,118	5,292
Poland	0	625	9,764	10,389
South Africa	0	2,010	641	2,651
Spain	0	0	6,290	6,290
Sweden	0	0	15,068	15,068
Tanzania	0	8,790	285	9,075
Türkiye	0	840	633	1,473
UK	0	0	5,368	5,368
USA	0	0	38,312	38,312
Total	550	60,202	142,145	202,897

Table 3. Means by Survey Mode

	CATI15	CAWI15	CAWI7
Flourishing	7.31	6.83	7.05
Life satisfaction	6.77	6.82	6.88
Happiness	7.01	6.99	6.97
Worthwhile	7.52	7.17	7.22
Physical health	7.58	7.04	6.68
Mental health	8.32	7.29	7.15
Purpose	8.24	7.19	6.95
Promote	8.28	7.61	7.77
Give up	8.00	6.90	7.26
Content	8.03	7.26	7.41
Satisfied relationships	7.79	6.98	7.00
Expenses	5.08	5.29	6.45
Worry food	5.25	5.51	6.90
Belonging	8.09	7.29	7.42
Purpose	8.24	7.19	6.95
Grateful	8.05	7.42	7.81
Trouble	7.12	6.87	7.98
Not lonely	6.58	6.45	6.77

Table 4. Random Effects Meta Analysis of Selected Variables by Survey Mode

a) Gender

		Survey Mode			
Dependent Variable / Gender		All	CATI15	CAWI15	CAWI7
Flourishing	Male	7.19	7.32	7.28	7.07
	Female	7.12	7.36	7.23	6.94
	Other	6.32	6.55	6.37	6.08
Not flourishing	Male	7.53	7.43	8.24	7.15
	Female	7.95	7.81	8.56	7.52
	Other	9.26	8.40	9.01	9.13
Life satisfaction	Male	6.82	6.90	6.74	6.89
	Female	6.89	7.10	6.91	6.82
	Other	6.00	6.32	6.30	5.69
Happiness	Male	6.99	7.04	6.93	6.99
	Female	7.03	7.19	7.15	6.91
	Other	5.96	5.99	6.43	5.68
No interest	Male	2.00	1.97	2.27	1.86
	Female	2.04	2.04	2.27	1.88
	Other	2.47	2.85	2.53	2.27

b) Employee and unemployed

		Survey Mode			
Dependent Variable / Employment		All	CATI15	CAWI15	CAWI7
Flourishing	Employee	7.20	7.40	7.34	6.99
	Unemployed	6.50	7.11	6.70	5.85
Not flourishing	Employee	7.70	7.45	8.16	7.36
	Unemployed	8.73	8.31	9.16	9.02
Life satisfaction	Employee	6.89	7.06	6.94	6.85
	Unemployed	5.97	6.60	6.02	5.38
Happiness	Employee	7.04	7.16	7.12	6.95
	Unemployed	6.30	6.81	6.38	5.73
No interest	Employee	2.01	1.95	2.24	1.87
	Unemployed	2.24	2.18	2.38	2.26

Table 5. Country ranks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Argentina	7	16	7	7	6	8	4	9	6	14	4
Australia	18	7	19	8	14	7	15	18	18	15	21
Brazil	13	21	10	9	7	11	5	10	21	21	7
Egypt	16	20	8	1	3	21	21	15	9	22	1
Germany	19	4	18	20	11	9	14	20	16	10	18
Hong Kong	9	13	13	15	9	6	8	12	12	8	8
India	15	15	12	4	10	18	19	19	10	7	22
Indonesia	1	2	1	2	1	4	1	5	19	6	9
Israel	2	4	2	11	5	1	3	4	3	20	4
Japan	22	5	22	22	19	15	20	8	5	4	20
Kenya	5	14	6	10	20	19	7	4	2	4	13
Mexico	3	8	3	4	2	4	2	7	7	19	3
Nigeria	4	18	4	16	15	16	9	3	22	12	5
Philippines	8	22	4	6	4	13	6	22	8	1	19
Poland	6	1	5	12	4	3	4	1	1	5	10
South Africa	10	10	11	13	18	14	12	13	14	18	2
Spain	17	11	15	14	13	10	13	11	20	2	11
Sweden	12	3	16	3	8	2	10	16	15	11	14
Tanzania	14	12	9	5	21	22	18	6	13	3	15
Türkiye	20	19	21	21	22	20	22	2	4	13	6
UK	21	9	20	18	16	12	16	21	11	16	16
USA	11	6	14	19	12	5	11	14	4	9	12
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	
Argentina	11	19	9	3	14	12	10	15	15	4	
Australia	21	4	19	4	4	4	19	14	19	9	
Brazil	20	12	10	13	15	14	11	9	9	3	
Egypt	22	3	14	16	22	22	9	2	12	2	
Germany	6	10	18	6	6	11	16	12	20	16	
Hong Kong	13	6	11	20	11	4	14	3	13	5	
India	10	9	12	19	20	6	7	6	6	13	

	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
Indonesia	5	1	1	11	2	2	1	5	2	18
Israel	1	4	3	1	13	13	4	1	11	6
Japan	4	22	22	22	21	20	22	19	22	22
Kenya	7	14	4	15	4	3	2	18	4	4
Mexico	4	16	4	2	16	10	4	13	10	8
Nigeria	9	8	2	18	1	1	5	4	1	1
Philippines	14	21	6	5	8	8	3	21	4	7
Poland	3	18	7	9	19	21	12	10	8	21
South Africa	2	20	8	14	3	9	6	4	7	12
Spain	15	11	13	7	5	15	13	22	16	15
Sweden	8	5	20	4	12	18	20	8	21	19
Tanzania	12	15	5	12	9	19	8	11	3	20
Türkiye	18	13	16	21	7	16	15	20	5	10
UK	19	2	21	10	10	5	21	16	18	14
USA	16	7	15	8	18	7	18	7	14	11

(1) GFS summed (2) Not flourishing (3) Flourishing (4) Belonging (5) Life satisfaction (6) Cantril (7) Happiness (8) Health problems (9) Gratitude (10) Balance of life (11) Mastery (12) Pain (13) Charity (14) Physical health (15) In trouble (16) Community (17) Volunteer (18) Purpose & meaning (19) Inner peace (20) Religion (21) Stranger.

Table 6. Not flourishing for 15 countries with CATI and CAWI

a) Not flourishing	CATI	CATI	CAWI	CAWI
25-34	.2614 (6.82)	.2419 (6.35)	-.3636 (8.43)	-.3184 (7.49)
35-44	.4392 (10.82)	.4027 (9.94)	-.8229 (18.48)	-.7288 (16.49)
45-54	.4193 (9.22)	.3713 (8.17)	-1.1439 (23.99)	-1.0250 (21.52)
55-64	.2870 (5.58)	.2544 (4.84)	-1.5440 (29.22)	-1.3649 (25.82)
65-74	.2410 (3.59)	.2377 (3.54)	-2.3946 (42.41)	-2.0296 (35.70)
75-84	.3267 (3.02)	.3270 (3.01)	-2.6181 (28.56)	-2.2225 (24.33)
85+	.1369 (0.51)	.2481 (0.93)	-2.3047 (7.30)	-1.9270 (6.19)
Female	.3641 (14.06)	.3654 (14.14)	.4666 (17.58)	.3652 (13.94)
Other sex	1.6850 (1.95)	1.9253 (2.17)	.7313 (3.75)	.3836 (1.95)
Empt & educn	No	Yes	No	Yes
Childhood	No	Yes	No	Yes
cons	6.8716	6.2388	9.2332	8.0739
Adj R ²	.0654	.0819	.1729	.2402
N	59,299	58,961	59,803	56,869
b) Flourishing	CATI	CATI	CAWI	CAWI
25-34	-.2046 (10.29)	-.1847 (9.40)	-.0116 (0.53)	-.0133 (0.61)
35-44	-.3326 (15.78)	-.2931 (14.02)	.0305 (1.36)	.0285 (1.27)
45-54	-.4179 (17.71)	-.3705 (15.79)	.0859 (3.56)	.0852 (3.53)
55-64	-.3982 (14.61)	-.3495 (12.89)	.3022 (11.30)	.3020 (11.30)
65-74	-.3878 (11.09)	-.3503 (10.08)	.9947 (34.79)	.9942 (34.80)
75-84	-.3802 (6.72)	-.3365 (5.96)	1.3113 (28.30)	1.3107 (28.29)
85+	-.4385 (3.15)	-.5167 (3.74)	1.1934 (7.35)	1.1835 (7.32)
Female	.0220 (1.64)	.0292 (2.19)	-.0293 (2.19)	-.0294 (2.19)
Other sex	-.6480 (1.45)	-.7152 (1.57)	-.5870 (3.92)	-.6106 (6.15)
Empt & educn	No	Yes	No	Yes
Childhood	No	Yes	No	Yes
cons	7.7290	8.1322	6.9076	6.9090
Adj R ²	.0574	.0857	.2299	.2307
N	58642	58,300	58,282	58,430

All models use country fixed effects. Employment and education include 3 education and 7 employment status controls including DK. Childhood controls include controls for health growing up (5), mother relation growing up (5), father relation growing up (5), parents marital status growing up (5), family income at 12 (4), outsider (1), attend religious services at 12 (4) with number of controls in parentheses. Argentina excluded. Omits abused as not asked in Israel.

Table 7. Seven countries CAWI only

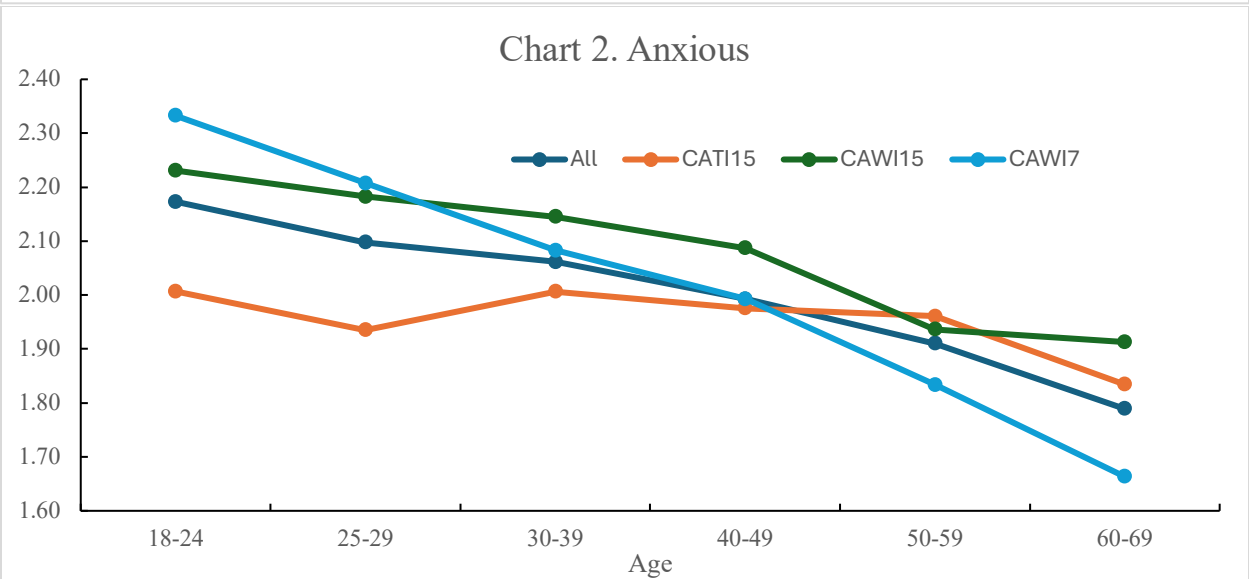
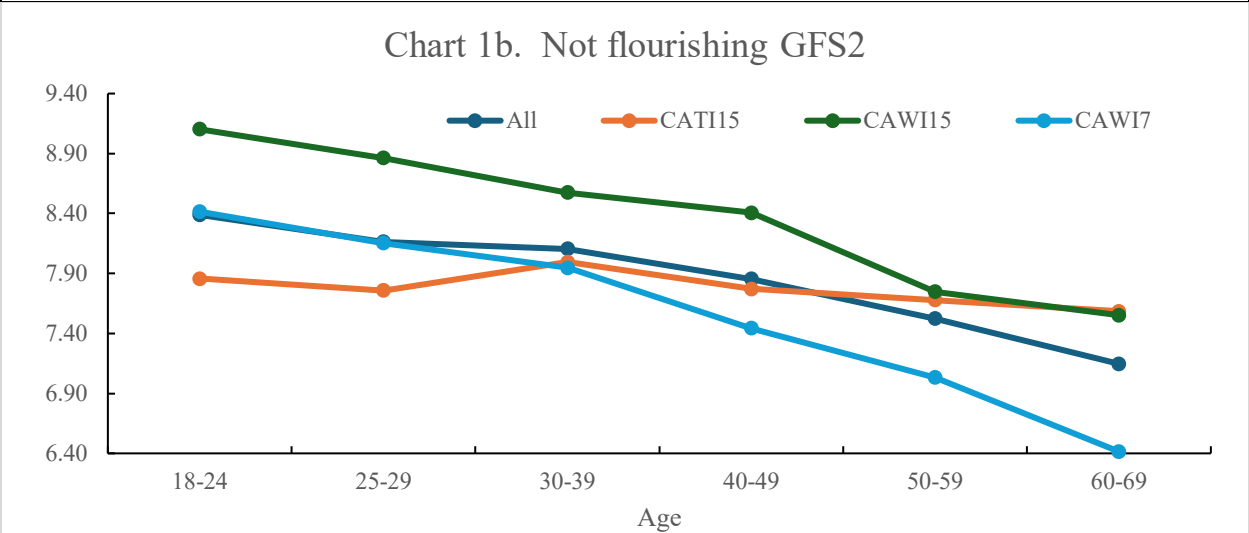
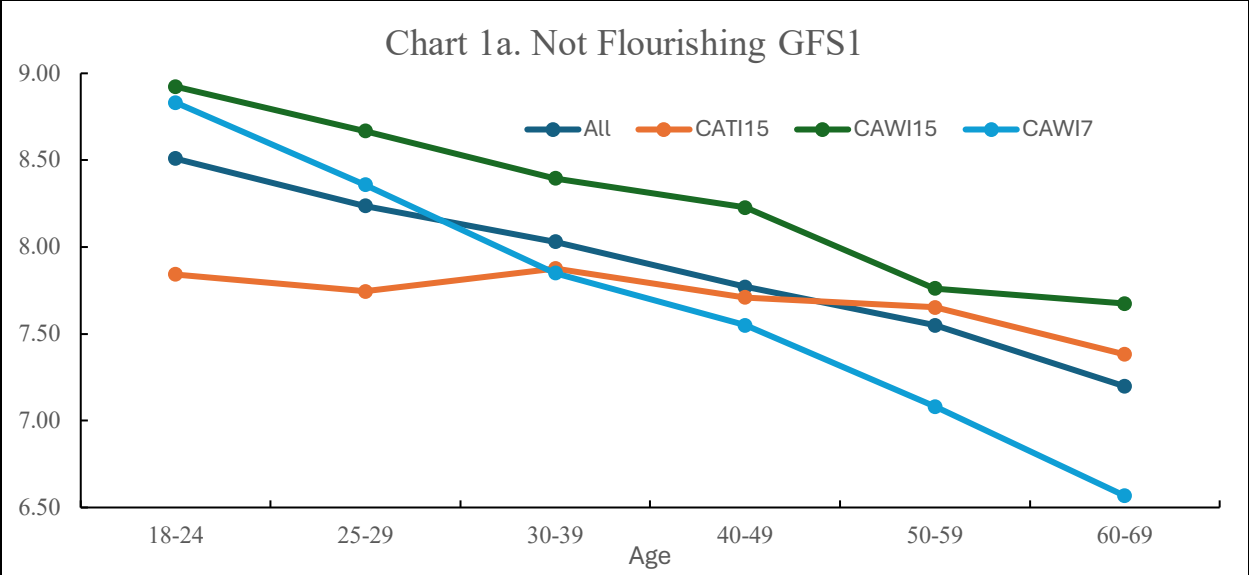
	Not Flourishing		Flourishing	
25-34	-.7094 (13.53)	-.7014 (13.74)	.1096 (4.02)	.1098 (4.31)
35-44	-1.1063 (21.36)	-1.0755 (21.25)	.1224 (4.55)	.1231 (4.87)
45-54	-1.5544 (30.07)	-1.4456 (28.49)	.2730 (10.16)	.2282 (9.01)
55-64	-1.9465 (38.46)	-1.7701 (35.46)	.5430 (20.63)	.4456 (17.88)
65-74	-2.4390 (47.89)	-2.1946 (43.48)	.8970 (33.87)	.6698 (20.96)
75-84	-2.6320 (47.26)	-2.3175 (42.04)	1.1598 (40.03)	.8711 (25.00)
85+	-2.5858 (26.91)	-2.2279 (23.77)	1.2512 (24.98)	1.0182 (21.70)
Female	.4104 (20.55)	.2609 (13.31)	-.1054 (10.19)	-.0121 (1.24)
Other sex	1.8981 (10.97)	1.2108 (7.50)	-.9938 (11.53)	-.5242 (6.52)
Empt & educn	No	Yes	No	Yes
Childhood	No	Yes	No	Yes
cons	8.3851	7.5887	6.7885	7.5638
Adj R ²	.1089	.1745	.1096	.2583
N	80,436	79347	79,484	78,418

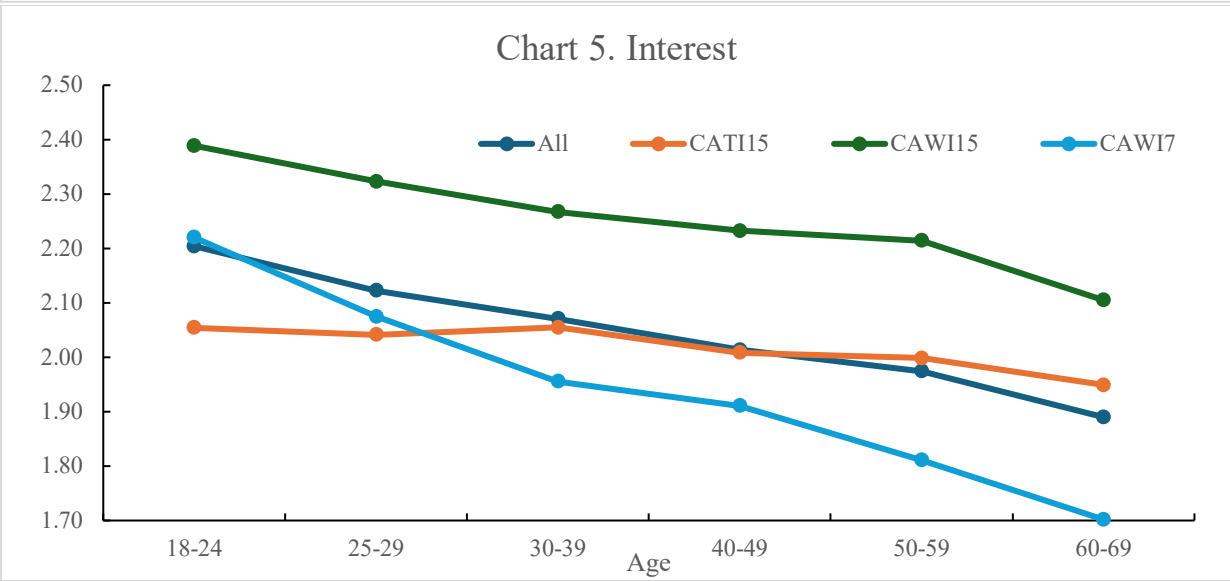
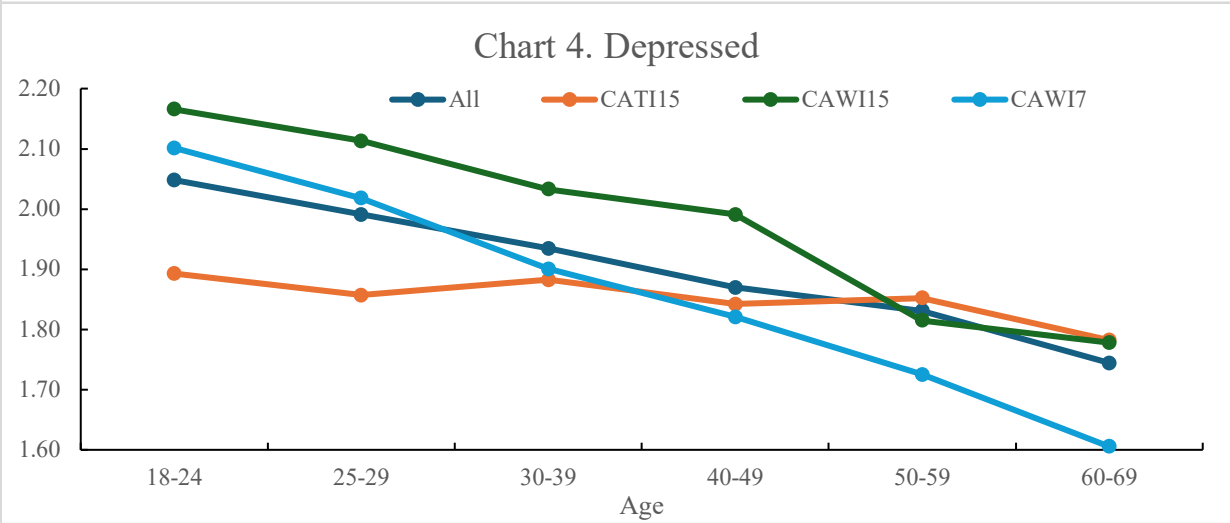
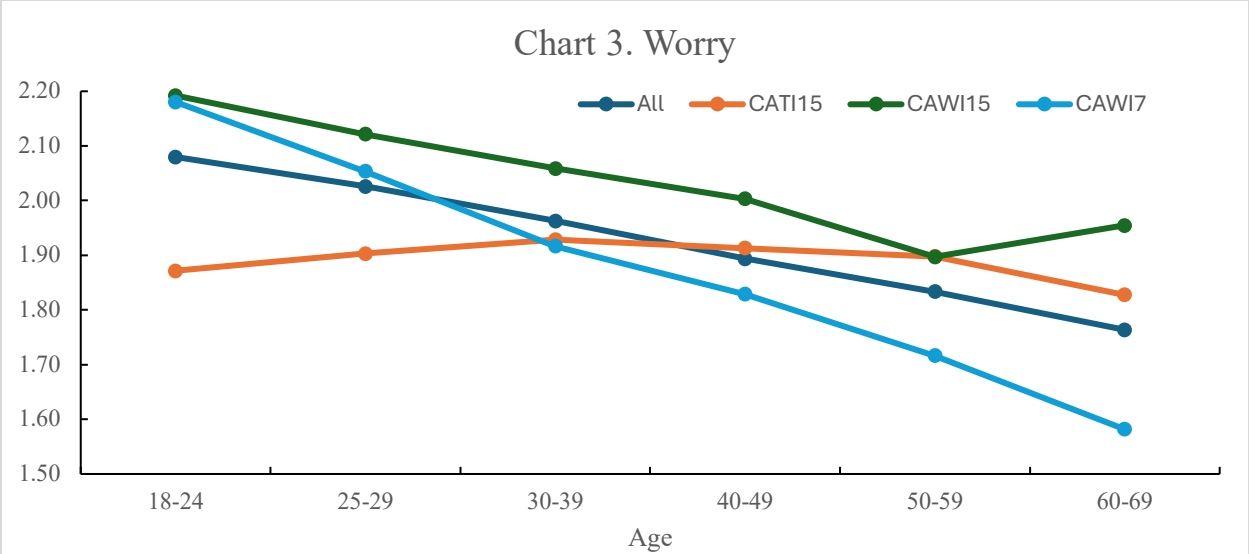
All models use country fixed effects. Employment and education include 3 education and 7 employment status controls including DK. Childhood controls include controls for health growing up (4), mother relation growing up (3), father relation growing up (3), parents marital status growing up (3), family income at 12 (3), abused (1), outsider (1), attend religious services at 12 (3) with number of controls in parentheses. Australia excluded. T-statistics in parentheses.

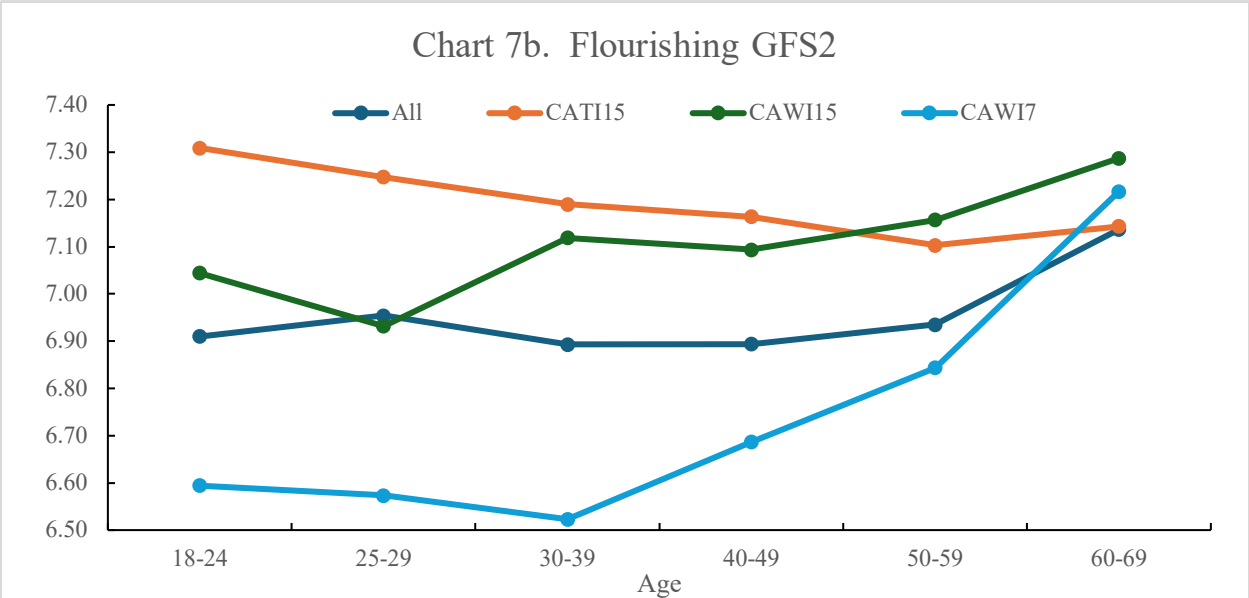
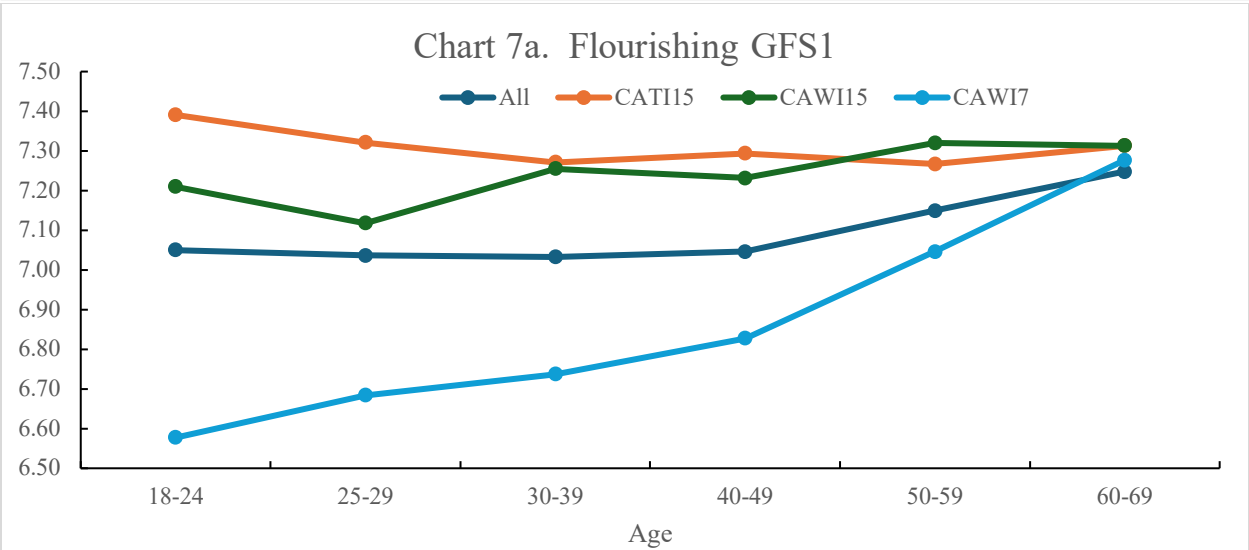
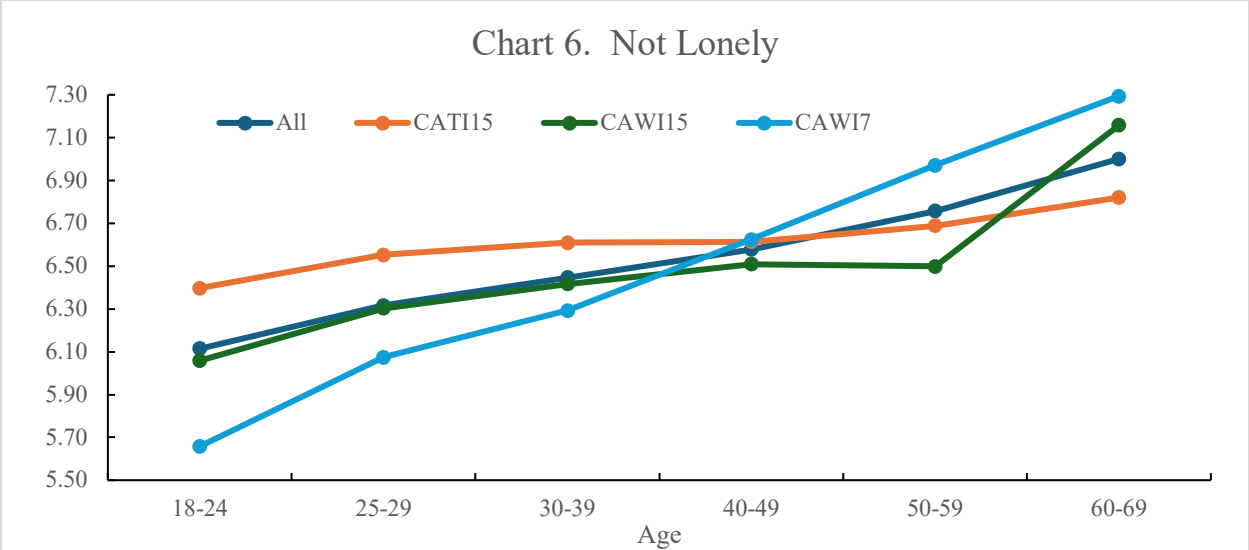
Table 8. Signs and significance of age 18-24 variable by survey mode

	CATI15	CAWI15	CAWI7
Flourishing	+	-	-
Not flourishing	-	+	+
Anxious	-	+	+
Depress	-	+	+
Worry	-	+	+
Interest	-	+	+
life sat	+	-	-
Happy	+	-	-
Life evaluation	+	+	+
Belonging	+	-	-
Life balance	+	-	-
Religious attend	-	-	+
Capable	+	-	-
Not lonely	+	-	-
Trouble	+	+	-
Community	+	+	+
Social support	+	0	0
Volunteer	-	+	+
Donate	-	-	-
Stranger	-	+	+
Grateful	0	-	-
Peace	0	-	-
Health problems	-	-	-
Pain	-	-	-
Mental health	+	-	-
Physical health	+	+	+
Worthwhile	+	-	-
Life purpose	+	-	-
Promote good	+	-	-
Give up	+	-	0
Content	+	-	-
Satisfied with relationships	+	-	-
Expenses	+	-	-
Worry safety	+	-	-

+ implies $t > 1.5$







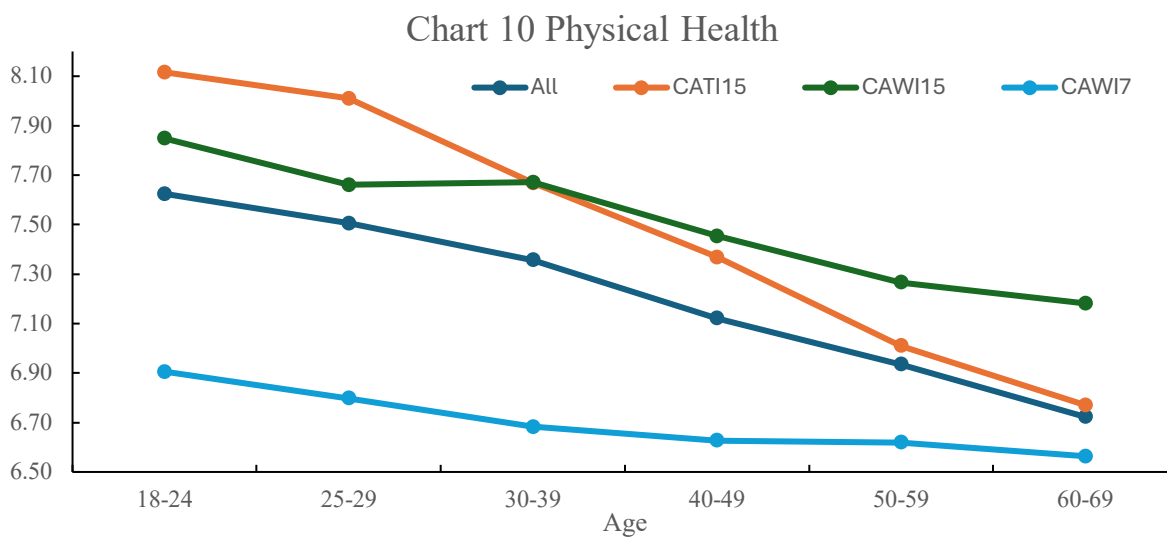
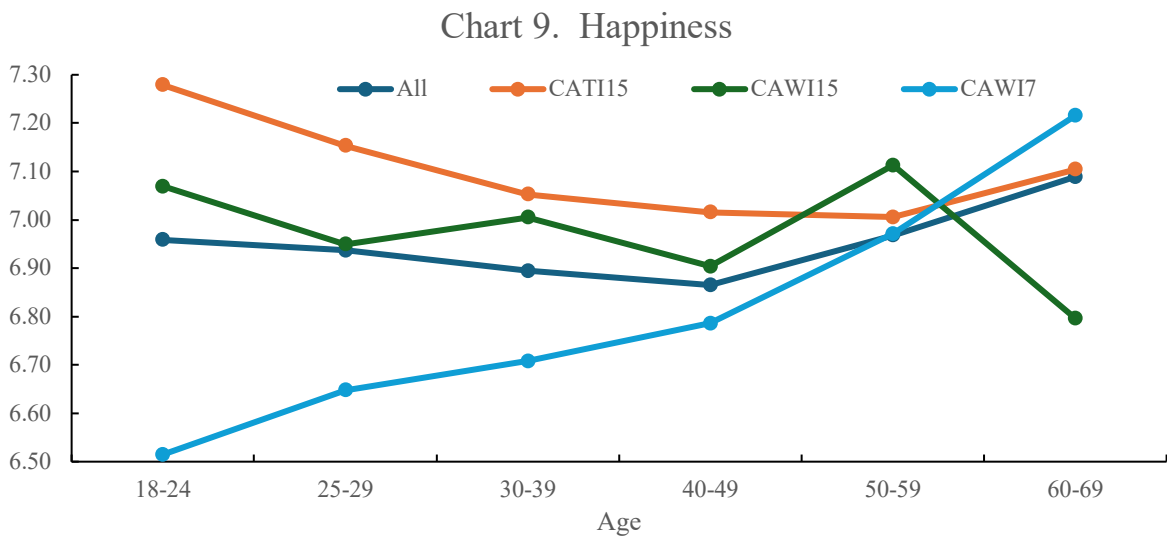
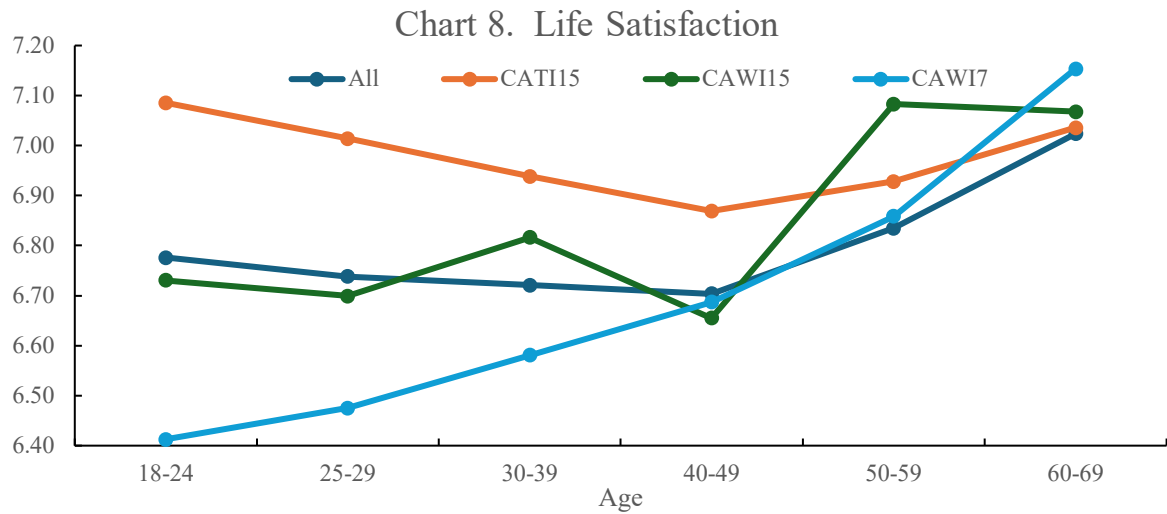


Chart 11. Mental health

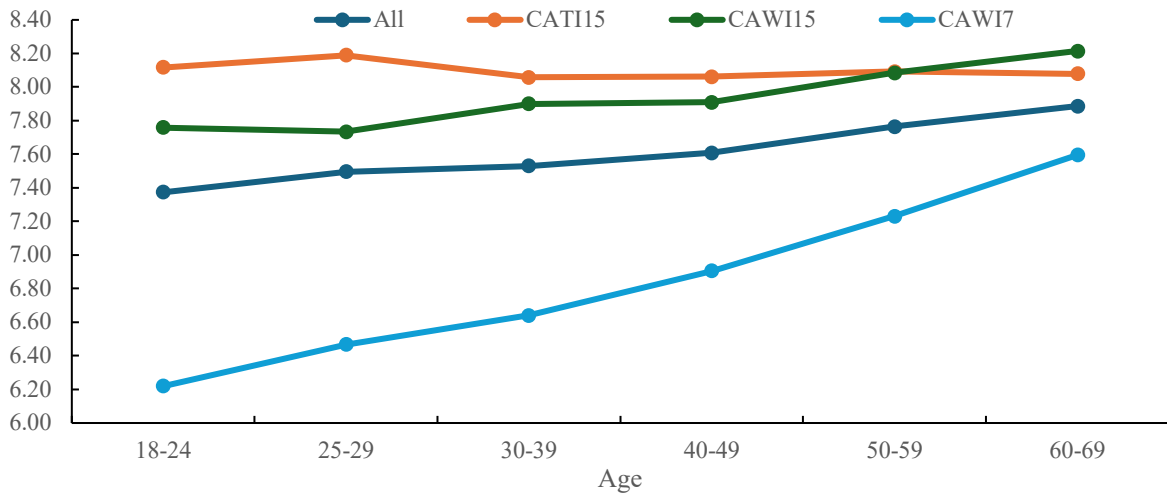


Chart 12. Worthwhile

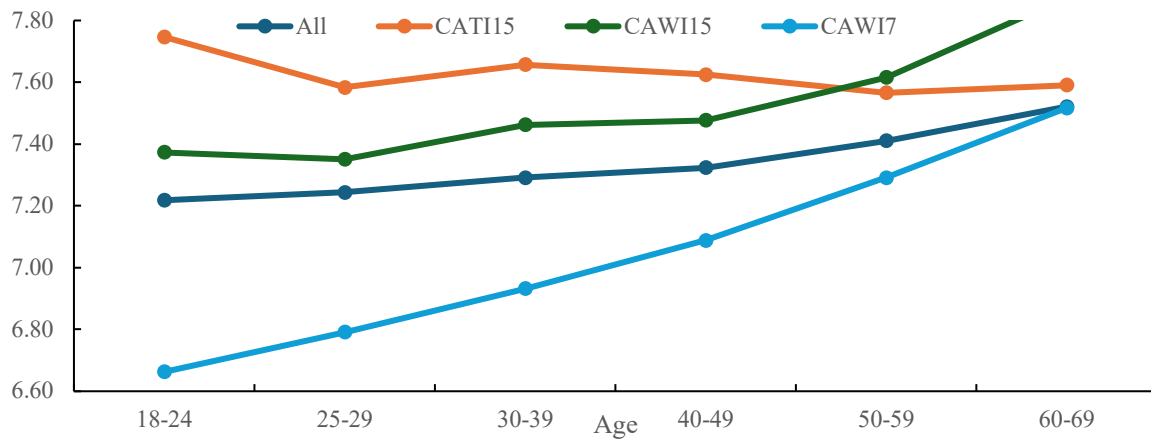


Chart 13. Life Purpose

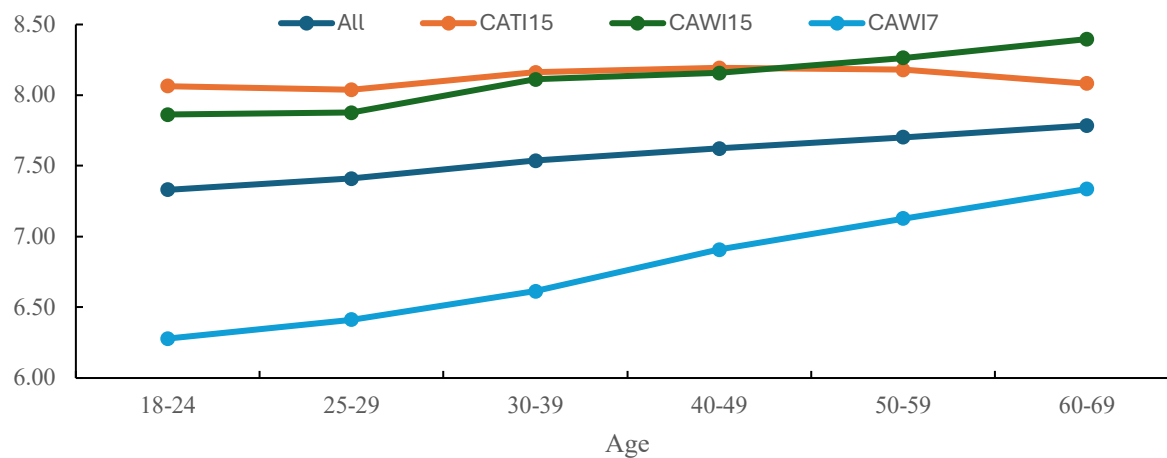


Chart 14. Promote good

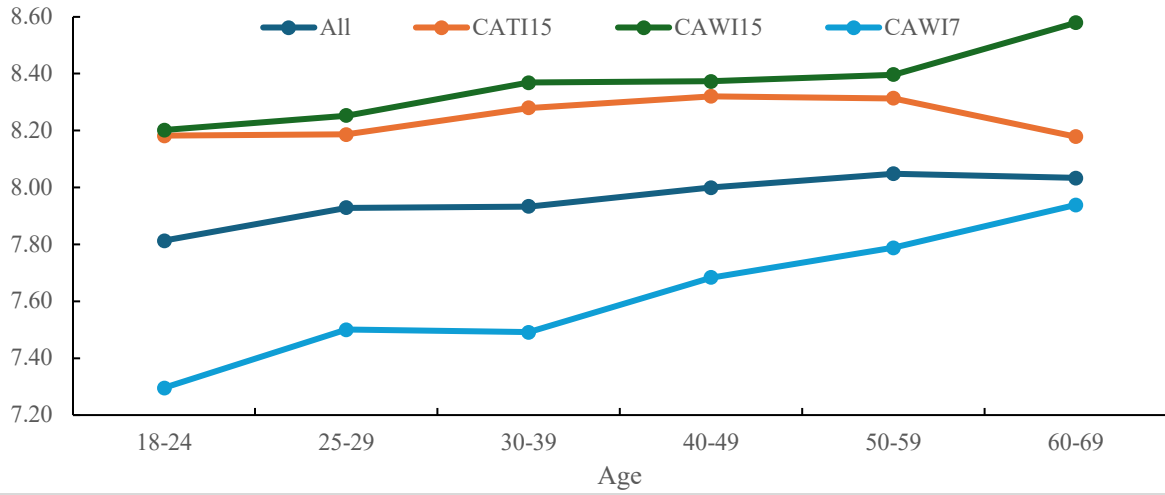


Chart 15. Give up

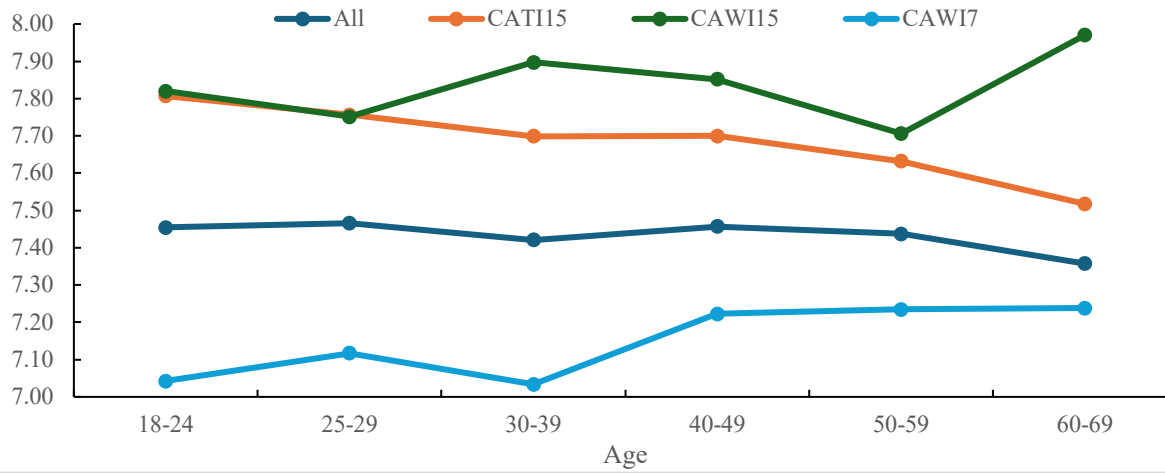


Chart 16. Content

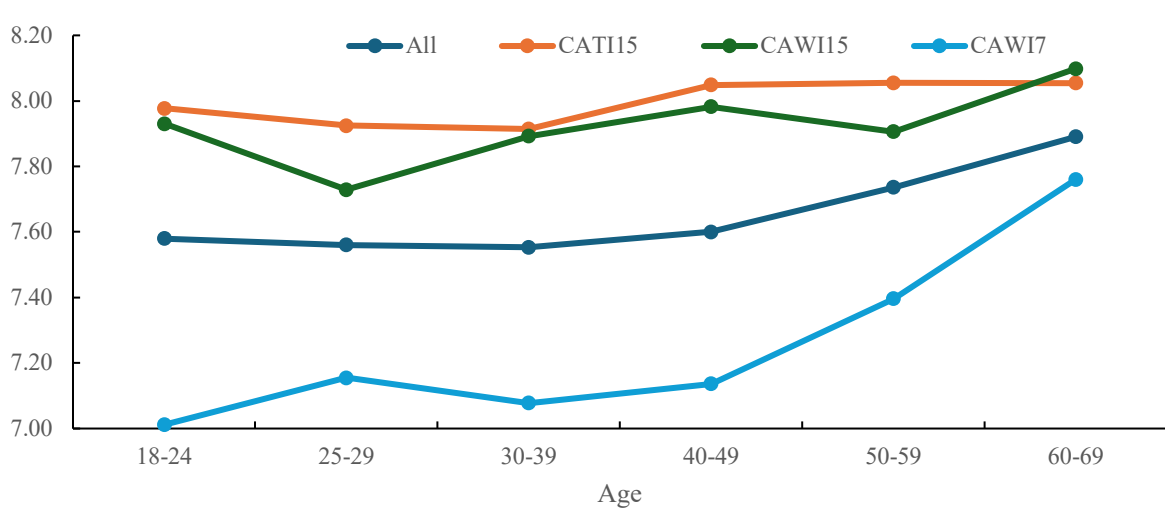


Chart 17. Satisfied with relationships

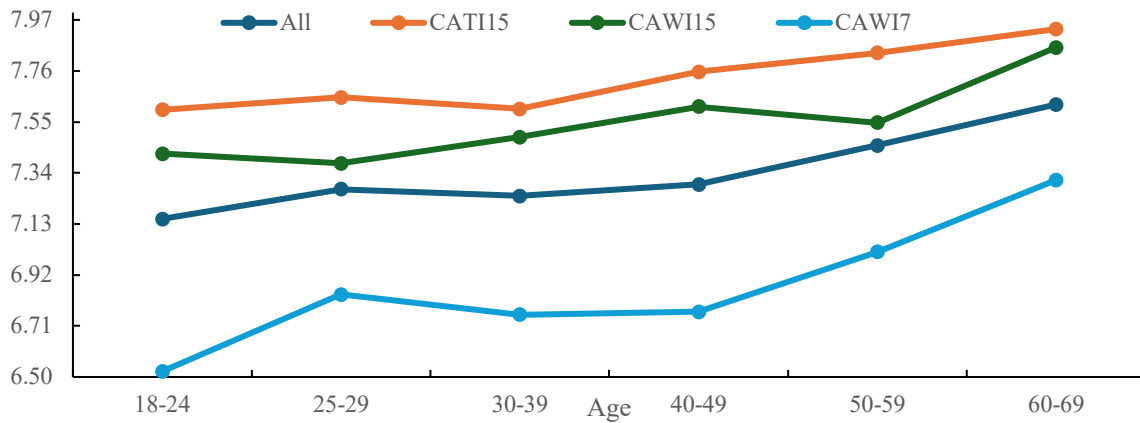


Chart 18. Expenses

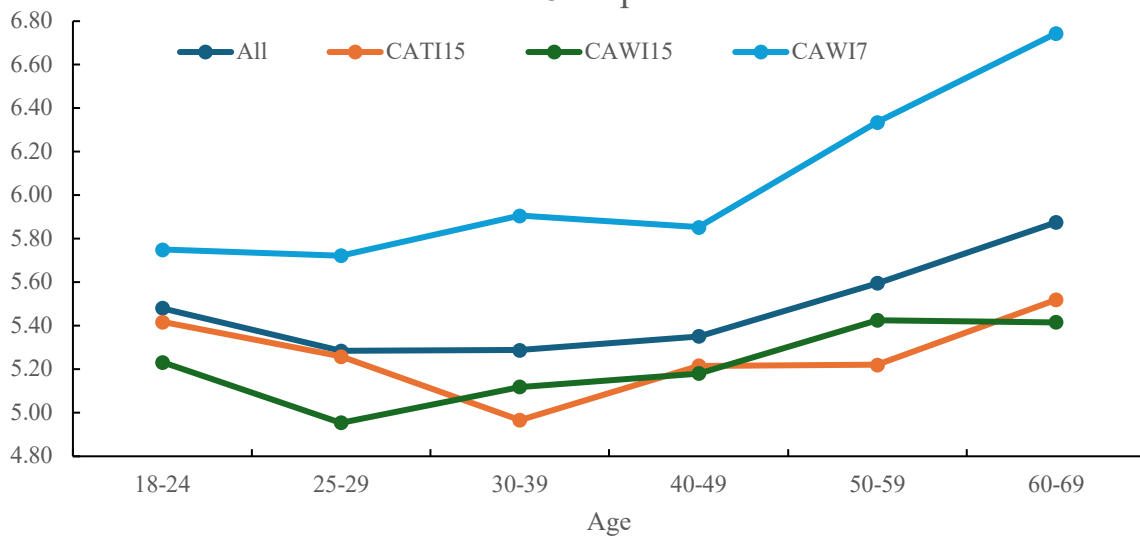


Chart 19. Worry About Safety, Food or housing

