Global Mind Project data in the United States: A comparison with national statistics

Joseph Taylor; Senior Researcher, Sapien Labs, 1201 Wilson Blvd, 27th floor, Arlington, VA, United States of America; joe@sapienlabs.org

Oleksii Sukhoi; Senior Statistician, Sapien Labs, 1201 Wilson Blvd, 27th floor, Arlington, VA, United States of America; oleksii@sapienlabs.org

Jennifer J Newson*, PhD; Senior Researcher, Sapien Labs, 1201 Wilson Blvd, 27th floor, Arlington, VA, United States of America; jennifer@sapienlabs.org

Tara Thiagarajan, PhD; Chief Scientist, Sapien Labs, 1201 Wilson Blvd, 27th floor, Arlington, VA, United States of America; tara@sapienlabs.org

* Corresponding author

Acknowledgements:

JT and TT developed the data acquisition methodology. OS carried out the analysis. JN and TT drafted the manuscript. All authors approved the final version. With thanks to the Sapien Labs team for assistance with data infrastructure. The full dataset from the Global Mind Project is freely available for not-for profit purposes from the Sapien Labs Researcher Hub. Access can be requested here: https://sapienlabs.org/global-mind-project/researcher-hub/

Funding: This work was supported by funding from Sapien Labs.

Key words: Population health; mental health; survey; methods; global mind project; MHQ; representativeness

Abstract

Population surveys have traditionally been conducted using mail, telephone, or face-to-face recruitment of randomly selected individuals within demographically stratified bands to acquire data samples that are representative of the population of interest. More recently, the growth of the internet has opened up the opportunity to generate large-scale samples at a much faster rate and lower cost. However, online participant recruitment methods typically result in non-probability samples that are subject to the black-box algorithms of online advertising platforms. This raises the question of whether it is possible to obtain data samples that demographically align with national statistics using web-based methods. Here we describe the online recruitment method used by the Global Mind Project (GMP) that serves advertisements through Facebook and Google Display inviting individuals to complete an anonymous, self-report mental wellbeing assessment which respondents take for the purpose of receiving a mental wellbeing score and personalized self-help report. Ads target respondents by age, sex and regional groups across a broad range of interest groups. Response rates are tracked by target group and dynamically adjusted to obtain the desired quotas. GMP data from the United States (83,589 responses from 2020 and 2023) were compared against time-matched data from the American Community Survey, Household Pulse Survey and the American Trends Panel for questions where there were exact or near-exact matches. These included educational attainment and marital status by age and biological sex, the percentage seeking treatment for mental health problems and number of close friends. Demographic trends match within 5-10% with a slight bias in GMP data towards single people with fewer friends who were seeking mental health treatment providing

a demonstration that GMP data obtained in the United States closely matches trends in census data.

Statement of significance:

This study demonstrates the feasibility of using a quota-based dynamic online ad targeting strategy for the collection of large-scale, anonymous data on mental wellbeing and its associated demographic, social, and behavioral trends. By comparing data from the Global Mind Project with established surveys from the United States, including the American Community Survey (ACS), Household Pulse Survey (HPS), and American Trends Panel (ATP), the findings highlight that the Global Mind Project's recruitment strategy generates results that closely reflect those from rigorous probability sampling methods. This alignment underscores the potential of this scalable, cost-effective, and globally adaptable recruitment strategy, and illustrates how it can mitigate some of the biases often associated with internet-based surveys, enabling rapid and real-time data collection that's critical for addressing the global mental health crisis.

Introduction

Traditionally, population surveys have been conducted using mail, telephone or face to face recruitment of individuals randomly selected within demographically stratified bands, typically by age, sex, geography and socioeconomics (Banerjee and Chaudhury 2010; Levy and Lemeshow 2013). This probability-sampling approach is considered the gold standard for acquiring data samples that are representative of the whole population of interest. However, collecting rigorous stratified samples using telephone or address-based recruitment methods is expensive, time consuming and difficult to scale globally, while non-compulsory surveys often struggle with low response rates. Furthermore, when asking about potentially sensitive or stigmatizing issues, such as those relating mental health, anonymity is needed to address concerns over data privacy or fear of self-disclosure.

Over the past few decades, the growth of the internet and mobile phone usage across the world (Data Reportal 2023) provides an opportunity for a new sampling and recruitment paradigm in the context of mental health research that can reach a large-scale and broad cross-section of the global population at a much faster rate and lower cost. However, online recruitment methods such as advertisements on Google or Facebook rely on non-probability sampling approaches that utilize the black-box optimization algorithms of the advertising platform, can potentially overrepresent certain populations or interest groups, and can only reach those who are on the chosen recruitment channel or who have a reliable internet connection, leading to a considerable debate as to the quality of data that they generate (Baker et al. 2013; Birnbaum 2004; Cornesse and Blom 2020; Couper 2000, 2007; Dutwin and Buskirk 2017; Fricker 2017; Goel et al. 2016; Kennedy et al. 2016; Schneider and Harknett 2022; Thornton et al. 2016; Whitaker et al. 2017). Furthermore, when online surveys are conducted anonymously, it leads to concerns about fraudulent, misrepresentative or bot responses within the sample (Glazer et al. 2021; Wang et al. 2023).

The Global Mind Project (GMP) uses online population sampling with the goal of providing a real time view of evolving mental wellbeing and the social, technological and lifestyle factors that drive it. It surveys various demographic, cultural, lifestyle, and life experience factors, including age, sex, ethnicity, education, employment status, and income, as well as 47 aspects of mental feeling and function on a 9-point scale, in 17 languages across 70 countries and has obtained data from 1.9 million people since its launch in 2020(Newson et al. 2022; Newson and Thiagarajan 2020). Individuals are anonymously recruited using paid advertisements placed on internet and social media channels including Facebook, Google Display and Google AdSense inviting them to complete a 15-minute online assessment. Rather than a asking individuals to participate in support of the research goals, the assessment provides individuals a free personalized report that synthesizes across their responses to provide mental wellbeing scores and self-help guidance. In addition, it uses a dynamically adjustable quota-based recruitment strategy (hereafter called Quota-based Dynamic Online Ad Targeting or Q-DOAT) which systematically targets pre-defined age-sex groups across a series of selected geographies using broad range of interest criteria and keywords with the goal of robust representation of the general population in each age-sex band for different countries of interest. Since it is not possible to pre-emptively seek fully proportionate representation of these demographic groups, the outreach aims to build sufficient representation within each demographic group such that representative outcomes can be obtained through post-stratification weighting (Goel et al. 2016; Pedersen and Kurz 2016).

Here we evaluated how demographic and social trends in GMP data for the United States (US) acquired using Q-DOAT compared with time-aligned trends from 3 national surveys conducted in the US: the American Community Survey [ACS; (US Census Bureau 2023a)] and Household Pulse Survey [HPS; (US Census Bureau 2024)] conducted by the US Census Bureau and the American Trends Panel (ATP) from the Pew Research Foundation. Comparisons were performed for questions where there were exact or near-exact matches. This included educational attainment and marital status by age and biological sex (our target criteria), the percentage seeking treatment for mental health problems, and number of close friends. It was hypothesized that there would be broad alignment of demographic trends, while social and behavioral responses would differ more due to differences in nonresponse bias across surveys.

Methods

Participant recruitment using Q-DOAT

Recruitment of participants in the US was conducted via English (from 2020 onwards) and Spanish language (from 2021 onwards) campaigns on Facebook, Google Display and Google AdSense with an advertisement containing the copy 'Get your mental wellbeing score: Fast, Free, Anonymous' along with a button linking to the start of the open survey. Anywhere from 30 to 100 advertisements were running at any given time in the US alone, that were regionally targeted towards each age-sex groups between 18 and 85 years, with advertisement spending in Spanish in proportion to the Spanish speaking population starting in June 2021. However, respondents to advertisements placed in any language had the option to change language on the assessment and respond in any of the languages offered which differed by year (only English in 2020, 3 languages in 2021, 9 in 2022 and to 14 languages in 2023).

The advertisements each used a broad range of interest keywords including self-awareness, self-development, health, wellness and coaching but no words specific to mental health or disorders in order to limit bias towards those with mental health problems. Such keywords are necessary to ensure relevance as per algorithms of Meta and Google which then seek 'look alike' audiences to optimize towards completion rates. Thus, such online advertisement targeting requires a trade-off between the breadth of interest of the audience and cost of completion, which is lower the more specifically advertisements are targeted.

Given this targeting paradigm, starts and completions were tracked for each advertisement within each source (Google and Facebook) using Urchin Tracking Module (UTM) codes and Google and Facebook Analytics, and advertisement spends were dynamically adjusted based on the demographic composition of respondents to further ensure sufficient representation across age, biological sex, and regional groups. Data from all new advertisements or sources were analyzed for parity before a new advertisement or source was scaled and included. In 2020, targets were set at a minimum of 500 responses for each age-sex group with at least 100 from each regional grouping (Northeast, Southeast, Southwest, Midwest and West) which would allow for post-stratification demographic weighting. This was increased to 1000 per group in 2021, with at least 150 from each regional grouping. Sample sizes against targets for age-sex groups for each year are shown in Table 1 and for each age-regional group in Table 2.

Table 1: Number of raw and clean US records for each age-sex group for each year

	Total records				Clean Records			
DEMOGRAPHIC	2020	2021	2022	2023*	2020	2021	2022	2023*
18-24 F	1776	2707	2761	1705	1641	2426	2482	1536
18-24 M	927	1730	1406	1139	857	1548	1286	1056
25-34 F	666	1434	1285	917	620	1326	1202	852
25-34 M	435	950	789	679	413	894	745	638
35-44 F	612	1139	1029	904	580	1079	969	828
35-44 M	388	677	657	635	366	629	624	593
45-54 F	927	1469	1213	996	890	1375	1141	943
45-54 M	552	885	881	735	527	840	833	692
55-64 F	1641	2179	2203	1847	1594	2098	2108	1764
55-64 M	1139	1839	1968	1790	1103	1777	1891	1719
65-75 F	2146	2955	3238	3282	2090	2869	3134	3162
65-75 M	1418	2757	2773	3135	1391	2673	2681	3013
75+ F	1163	2327	2639	3777	1141	2257	2550	3651
75+M	598	1449	1898	2764	577	1414	1842	2659
TOTAL	14388	24497	24740	24305	13790	23205	23488	23106
			TOTAL	87930			TOTAL	83589

*partial year; data recorded for this analysis ended in November 2023

Table 2: Number of clean records collected in 2022 for each age-sex group for each ES region

Region	18-24	25-34	35-44	45-54	55-64	65-74	75+
Northeast	613	373	277	372	760	1160	795
Southeast	809	433	352	519	1125	1703	1293
Southwest	450	218	177	188	406	507	429
Midwest	809	425	397	478	934	1367	982
West	1083	494	386	413	774	1075	893
TOTAL	3764	1943	1589	1970	3999	5812	4392

We note that costs per start and complete can vary considerably across demographic groups and are particularly expensive for males between 25-44 in the US resulting in clean record numbers slightly below target for this demographic. On the other hand, numbers for females and older adults can easily exceed targets even at minimum budgets. Budgets permitting, this method allows for even more granular targeting and dynamic balancing than used here.

The language, recruitment methods and assessment were approved by the Health Media Lab Institutional Review Board (HML IRB; OHRP Institutional Review Board #00001211, Federal Wide Assurance #00001102, IORG #0000850). Participants took part in the online survey voluntarily, anonymously, and without any financial compensation. Participants consented to take part by clicking on a start button after reading a detailed privacy policy.

GMP Data Processing and Quality Checks

While bots are unlikely to be a challenge as they will not typically be served advertisements, one of the challenges of anonymous online recruitment is that people may simply click through the survey to view its content. In order to eliminate responses of this nature, multiple data cleaning criteria were applied. First, responses that were completed in under 7 minutes (the minimum time needed to read all questions) or over 60 were excluded from the analysis. Second, those who had responses with a standard deviation of less than 0.2, representing people who answered with the same rating value across all 47 rating items (e.g. 111… or 5555…), were excluded. Finally, records where the respondent answered ‘No’ to the question ‘Did you find this assessment easy to understand?’ were removed. Completions arising from organic traffic, including peer referrals, were also excluded as they lay outside of the managed targeting criteria. This resulted in 5% of responses being excluded from the analysis. After cleaning, the data sample size was 13,790 in 2020, 23,205 in 2021, 23,488 in 2022 and 23,106 in 2023. However, N values for individual analyses below differed due to removal of blanks or ‘prefer not to say’ selections in the questions of comparison.

Comparison against the ACS, HPS and ATP Data

GMP US survey data was compared against data from the ACS, HPS and ATP. As the GMP collects data on a wide variety of demographics, cultural, lifestyle and life experience factors, only questions where there was an exact or near-exact match to questions in the ACS, HPS and ATP were selected for inclusion in this study to ensure an accurate comparison. This included educational attainment and marital status by age and biological sex from GMP and ACS obtained in 2022; the percentage seeking treatment for mental health problems from GMP and HPS from 2020 to 2023; and number of close friends from GMP and ATP for 2023.

For each of these data elements within GMP, the percentage of respondents selecting each answer option were computed for each age and biological sex group. In all cases, this was done by first computing averages for each age-sex group and then computing a weighted average based on the proportion of the population in each age group as provided by the United Nations (UN) Population Statistics (United Nations 2022). For each of the comparison surveys (ACS, HPS, ATP), precalculated numbers by age and biological sex or national aggregates were directly downloaded from the respective survey sites.

Comparison of marital and education status data in GMP and ACS

ACS 2022 data was downloaded from the ACS data site (US Census Bureau 2023b). The specific ACS tables downloaded were S1501 (Educational Attainment; N~3 million) and B12002 (Marital Status; N=~3 million). ACS Table S1501 reports education attainment as percentage with High School or Higher and percentage with Bachelor's or Higher. To create a comparable metric from the equivalent 2022 GMP data (n=22,396), the percentage with

Bachelor's degree, Master's Degree and PhD degree were summed as percentage Bachelor's or higher, while percentage with High School and Associate degree were added to this to arrive at percentage High School or Higher.

ACS marital status options provided were: Never Married, Married/Spouse Present, Married/Separated, Married, Spouse absent, Divorced, Widowed while GMP marital status options provided were: Single (Never Married), In a relationship, Married/Civil Partnership, Divorced/Separated, Widowed, Prefer not to say (n=22,396). Given the slight differences, the data were aggregated and compared as follows: (i) ACS Never Married to GMP Single (Never Married) + In a relationship; (ii) ACS Divorced + Married but Separated to GMP Divorced/Separated; (iii) ACS Married/Spouse present to GMP Married/Civil Partnership; (iv) ACS Widowed to GMP Widowed.

Comparison of mental health treatment status in GMP and HPS

The percentage seeking treatment for mental health problems captured as part of the HPS from January 2020 to October 2023 (N=2,036,992) was compared to the equivalent information captured by GMP across the same time period (N=70,800 after removal of those answering 'Prefer not to say'). The HPS asked the following questions:

HPS1: At any time in the last 4 weeks, did you take prescription medication to help you with any emotions or with your concentration, behavior, or mental health? Yes/No

HPS2: At any time in the last 4 weeks, did you receive counseling or therapy from a mental health professional such as a psychiatrist, psychologist, psychiatric nurse, or clinical social worker? Include counseling or therapy online or by phone. Yes/No

While GMP asked:

Project GMP1: Are you presently undergoing treatment for any mental health challenges? Yes/No/Prefer not to say

Specifically, the percentage of those who answered ‘Yes’ to GMP1 above was compared to the percentage who answered ‘Yes’ to either HPS1 or HPS2. Age-sex weighted national estimates from 2020 to 2023 were also compared, as well as the age-wise break-up for 2023 alone. Calculated results for those who answered ‘Yes’ to either HPS1 or HPS2 by age and year were downloaded directly from the CDC site (CDC 2022). Annual estimates were arrived at by averaging results across multiple time periods of the HPS survey reported during the same year. The HPS reported aggregated data by age bands 18-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80 years and above, while GMP captured age data in bands 18-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84. 85+. Since only pre-aggregated results were available for the HPS this did not afford a perfectly age-aligned comparison.

Comparison of number of close friends in GMP and ATP

The average percentage of the population with each number of close friends from 0 to 5+ in the ATP, obtained during July 2023 (N=5,057) was compared against the equivalent GMP data obtained between January 1st and November 30th 2023 (N=19,857 after removal of blanks and values over 100). The results from the ATP used responses to the question: ‘Not counting your family, how many close friends do you have?’ with answer options 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more. The average percentages of the population reporting each number of close friends in the ATP were weighted to be representative of the US adult population by sex, race, ethnicity, education, and other categories. Equivalent percentages were computed from GMP data for the similar question: ‘How many close friends do you have?’ with a numerical response field. These were age-sex weightings used national statistics from UN to reflect their proportions in the population.

Results

Demographic Trends in the ACS are mirrored in GMP data

Figure 1 shows a comparison of educational attainment and marital status by age group for data obtained by ACS (N=~3.5M) and GMP (N=22,396) in 2022. Overall, trends for both educational attainment (Figure 1A) and marital status (Figure 1B) in GMP data closely matched trends in the ACS data. However, a few differences were noted. For educational attainment, the proportion of High school or higher was 12% higher in GMP than ACS in the 18-24 age group. For other age groups, the difference ranged from ±0% to ±3% (average ±1%). GMP age question captures 18-24 as a group, therefore it is possible that this group either contained a smaller proportion of 18-year-olds, the majority of whom are still in high school, or that 18-year old's close to completing high school choose "High School" for educational attainment in preference over 'Some High School' which was the alternative option available. Presently, those still in high school are directed to a youth version of the survey. For marital status, GMP data showed a higher proportion of never married respondents across all ages except 18-24 (average difference 6%; range ±1% to ±9%) and a correspondingly lower proportion of married respondents (average difference 7% difference; range ±1% to ±10%). For 18–24-year-olds, there was a higher proportion of never married respondents in ACS data (7%), while the proportion married was similar (1% difference). 3% chose 'Prefer not to say' in GMP and were not included which may explain some of the difference between married and never married.

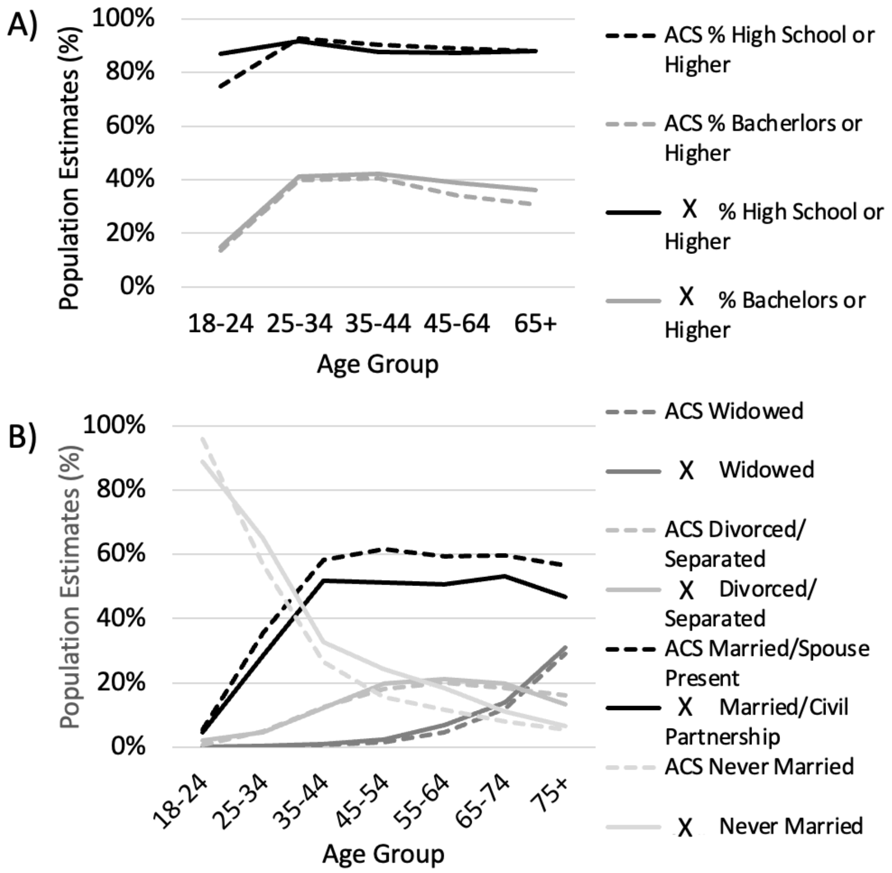

Figure 1. Comparison of population estimates (%) for educational attainment (A) and marital status (B) by age group for data obtained by ACS (N= ~3.5 M; dotted lines) and GMP (N=22,396; solid lines) in 2022.

Reported mental health treatment seeking behavior captured in the HPS was mirrored in GMP

Figure 2 compares trends of the percentage of the adult population seeking professional treatment for a mental health problem over time from 2020 to 2023, and by age for 2023, between the HPS and Project X. The specific questions asked by the HPS and GMP surveys were similar but not identical. While GMP asks about 'current' treatment for mental health challenges without specifying which type of treatment, HPS asks specifically whether prescription medication and/or therapy/counseling were taken in the past 4 weeks.

Nonetheless, it provides a broadly similar comparison that can determine if GMP oversamples for individuals with mental health problems. Figure 2A shows that the age-sex weighted national estimates of GMP were within ±1% of the national estimates of the HPS for all years other than 2021 where it was 5% higher. The estimates by age for 2023 data are shown for HPS and GMP in Figures 2B and 2C respectively. HPS data tables use different age categories (e.g. 30-39, 40-49 rather, than 35-44, 45-54 etc.) precluding a direct comparison. However, broadly, the percentage seeking treatment in GMP data was generally higher for ages 25-54 by an average of 8% (range 6% to 10%) and lower for ages 70 and above by an average of 5% (range 4% to 7%). This difference could arise because GMP responses consider any treatment beyond prescription medication and therapy/counseling. However, it may also arise from a nonresponse bias where younger people in treatment were more likely to take part in GMP while older people in treatment may be less likely to be on the Internet.

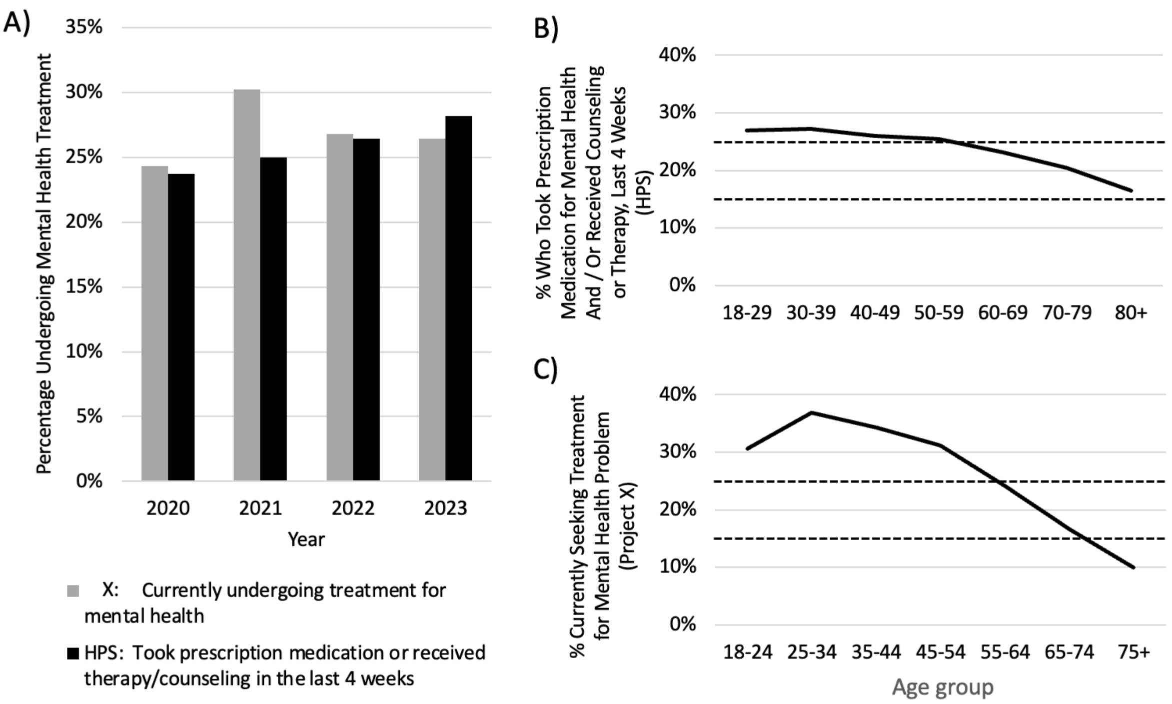

Figure 2. Comparison of trends for the percentage of the adult population seeking professional treatment for a mental health problem over time and by age between the HPS and GMP. (A) Comparison of the percentage currently undergoing treatment for mental health problem (GMP, grey) and the percentage who took prescription medication or received therapy/counseling in the last 4 weeks (HPS, black) from 2020 to 2023. (B) HPS data across different age groups for the percentage who took prescription medication for mental health conditions and/or received counseling or therapy in last 4 weeks in 2023. (C) GMP data across different age groups for the percentage currently seeking treatment for mental health problem in 2023.

National trends of close friendships in the ATP compared to GMP

Figure 3 shows a comparison of the number of close friends reported in the ATP in July 2023 compared to the annual GMP data for 2023. While the pattern across number of friends was similar, there were some key differences. Respondents in GMP sample were more likely to report only 2 close friends or less (average 15%) compared the ATP (average 10%; difference 5%) and correspondingly less likely to report 5+ friends (28%) compared to the ATP (38%; difference 10%). The differences may arise for multiple reasons including differences in the estimation methods or differences in the nature of nonresponse bias between the two surveys.

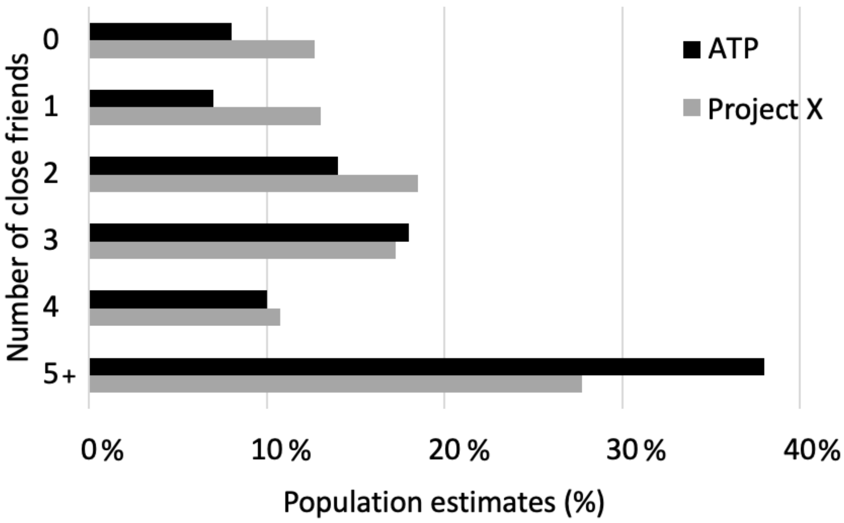

Figure 3: Comparison of population estimates (%) for the number of close friends (0 to 5+) reported in the ATP in July 2023 (black) compared to the annual GMP data for 2023 (grey).

Discussion

Principle results

We have shown here that US data, obtained by GMP using quota-based dynamic online ad targeting (Q-DOAT), closely aligns with national trends obtained from various rigorously stratified and randomly sampled US based surveys such as the ACS and HPS conducted by the US Census Bureau, and the ATP conducted by Pew Research Centre. This includes demographic factors of marital status and educational attainment, mental healthcare trends and friendship, which represent a diverse range of variables. Altogether, this suggests that data obtained anonymously through Q-DOAT, a dynamically responsive online recruitment method, aligns well with data obtained from identified participants recruited through rigorous probability sampling methods and can be reliably used to explore relationships between factors in the general population. This has particular importance given the challenges with many probability survey approaches [e.g. logistically complex, time intensive, expensive, increasing nonresponse rates, difficult to scale globally (Brick and Williams 2013; Keeter et al. 2006; Kohut et al. 2012; Leeper 2019; US Government Accountability Office 2002)] and the advantages of GMP recruitment methodology. In particular, GMP is able to rapidly recruit participants (1000-2000 people take part globally every day); is ten to twenty times more cost effective (average cost per respondent ranges from $0.15 to $3.5 depending on country, state, language and demographic group); has global scalability (it currently runs in 70 countries); is adaptive to changing societal trends and events; and is readily able to target specific populations of interest. Furthermore, when asking about potentially sensitive or stigmatizing issues, such as those relating mental health, anonymity is needed to address concerns over data privacy or fear of self-disclosure. It also positions GMP as an easily scalable and flexible platform for tracking national trends, and in particular emerging trends. The findings also contribute to the growing body of literature that highlights the opportunity of using online channels such as Facebook and Google Ads to recruit participants for health-related studies, especially when targeting is carefully considered and dynamically updated based on real-time demographic profiling of the data sample (Astley et al. 2021; Batterham 2014; Schneider and Harknett 2022; Thornton et al. 2016).

Nonresponse bias and Limitations

Understanding nonresponse bias is important as every survey will have its own form of nonresponse bias based on the survey topic and mode of delivery. GMP X aims to target a general population rather than specifically those who have mental health problems, but by necessity it uses targeting keywords that include terms such as self-awareness, health and wellness which raises the possibility of bias related to interest in these aspects. In this context, although GMP data closely mirrored national trends, it is important to note some differences arising between the datasets. In GMP data, we observed a bias towards single people (as compared to the ACS), more people seeking treatment in the age groups between

25-54 (as compared to the HPS) and a higher proportion with fewer friends (as compared to the ATP). These differences were on an order of magnitude of 5-7%. This suggests the possibility that those who were more inclined to take part in GMP were single people with fewer friends who were seeking mental health treatment or, conversely, a nonresponse bias towards married people with lots of friends who are not seeking mental health treatment (or both). This latter finding aligns with other studies involving mental health surveys that report a greater representation of people with mental health problems within the sample (Batterham 2014; Lee et al. 2020). However, other explanations are also possible. For example, GMP considers all treatments for mental health and not just prescription medication and therapy/counselling as included in the HPS. Some fraction of the difference may therefore reflect those undergoing other types of treatment (e.g. brain stimulation, neurofeedback). With respect to the differences in trends for the number of close friends, one explanation for the difference might be that people who frequent social media channels may have fewer close friends. Alternatively, as the ATP recruits participants for a broader survey of civic trends it is possible that those who agree to participate in the ATP are more civic minded which biases towards people with more friends. Nonetheless, while these differences could explain and mitigate some of the differences observed, a small nonresponse bias in GMP data towards those with greater risk factors for mental health challenges cannot be ruled out.

One of the primary goals of GMP is to track country level trends across the globe, particularly with respect to mental health status. The question therefore arises as to how much a 5-7% bias towards single people, those with few friends and/or those taking prescription medication or in therapy would shift these estimates. For example, the difference in the percentage of those who are mentally distressed or struggling [as measured by the MHQ score, see (Newson et al. 2024) for more details on the MHQ and how it is calculated] among those seeking treatment versus those not seeking treatment is only 14% (40% vs 26%). The average percentage of those distressed or struggling (MHQ scores<0) re-weighted by the proportions of those single and seeking treatment in each age group, as per the ACS and HPS respectively, results in a value 1-2% lower for most age groups, providing an estimate of the magnitude of this potential bias. Altogether, this suggests that national trends of the percentage distressed or struggling in GMP are overestimated by up to 2%, particularly for middle age groups. However, as the recruitment methods are relatively stable each year, changes over time would still provide a reliable estimate of the magnitude of change.

Dynamic vs Static strategies

It is important to point out, however, that these results do not mean that all internet-based surveys using online recruitment strategies can be assumed to be well aligned with nationally representative samples. The Q-DOAT method differs from typical static river sampling advertisement strategies in that it utilizes a complex and dynamic optimization of demographic targeting, key words and other factors to recruit a sufficiently broad-based sample. This method requires sophisticated real-time analytics of responses and frequent adjustment of targeting. The strategies of GMP have thus been optimized by multiple experiments and involve a large number of recruitment advertisements (currently 800+ globally) with diverse targeting that are actively managed to accomplish the results shown here. Many online studies still report biases, e.g. (Lee et al. 2020) and emphasize the need for careful targeting and advertisement creation. For example, if recruitment were carried out through advertisements served to people searching for information on mental health disorders, while response rates may be higher (Batterham 2014), the mental wellbeing profiles would skew towards worse mental health than the general population, and therefore the percentage seeking mental health treatment would be higher than the national metrics reported by the HPS.

GMP data beyond the United States

GMP presently operates in 17 languages across 70 countries although sample sizes vary across countries. While it is not possible to directly extrapolate these conclusions to all other countries, we note that the same methodology is used across the world, suggesting the potential for similar outcomes. However, it must be noted that GMPrecruits only from the Internet-enabled population. With 94% of the US population Internet-enabled, most of the population in the US are covered and may be invited to participate. In contrast, Project X data will increasingly deviate from national statistics with decreasing Internet penetration and, for countries with substantially lower internet penetration. In future, we will present results for other country datasets against nationally available statistics of internet-connected populations, also noting that there is currently very little comparative statistics on the online populations of non-western countries (Sanchez et al. 2020).

Conclusion

Here we have shown that GMP data obtained using the Q-DOAT method shows good alignment with large surveys using more rigorous sampling techniques, suggesting that GMP data from the US are likely to be broadly reflective of the national population and positioning the project as a rapidly scalable real-time view of mental health and wellbeing as well as demographic and social trends. Altogether, with rising rates of mental health conditions around the world, especially in younger populations (CDC 2023; Sapien Labs 2021; Twenge et al. 2019), there is an urgent need for a new paradigm of data collection within the field of mental health, something also noted by Sanchez and colleagues (Sanchez et al. 2020) who stated *"Developing new strategies to increase recruitment for mental health research is essential to addressing the field's most pressing problems.*

Conflicts of Interest:

None declared.

References:

Astley, C. M., Tuli, G., Mc Cord, K. A., Cohn, E. L., Rader, B., Varrelman, T. J., Chiu, S. L., Deng, X., Stewart, K., Farag, T. H., Barkume, K. M., LaRocca, S., Morris, K. A., Kreuter, F., and Brownstein, J. S. (2021), "Global monitoring of the impact of the COVID-19 pandemic through online surveys sampled from the Facebook user base," *Proceedings of the National Academy of Sciences of the United States of America*, 118, e2111455118. https://doi.org/10.1073/pnas.2111455118.

Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., Gile, K. J., and Tourangeau, R. (2013), "Summary Report of the AAPOR Task Force on Non-probability Sampling," *Journal of Survey Statistics and Methodology*, 1, 90–143. https://doi.org/10.1093/jssam/smt008.

Banerjee, A., and Chaudhury, S. (2010), "Statistics without tears: Populations and samples," *Industrial Psychiatry Journal*, 19, 60–65. https://doi.org/10.4103/0972-6748.77642.

Batterham, P. J. (2014), "Recruitment of mental health survey participants using Internet advertising: content, characteristics and cost effectiveness," *International Journal of Methods in Psychiatric Research*, 23, 184–191. https://doi.org/10.1002/mpr.1421.

Birnbaum, M. H. (2004), "Human Research and Data Collection via the Internet," *Annual Review of Psychology*, 55, 803–832. https://doi.org/10.1146/annurev.psych.55.090902.141601.

Brick, J. M., and Williams, D. (2013), "Explaining Rising Nonresponse Rates in Cross-Sectional Surveys," *The ANNALS of the American Academy of Political and Social Science*, SAGE Publications Inc, 645, 36–59. https://doi.org/10.1177/0002716212456834.

CDC (2022), "Mental Health Care: Household Pulse Survey."

CDC (2023), *Youth Risk Behavior Survey: Data Summary & Trends Report*.

Cornesse, C., and Blom, A. G. (2020), "Response Quality in Nonprobability and Probability-based Online Panels," *Sociological Methods & Research*, 004912412091494. https://doi.org/10.1177/0049124120914940.

Couper, M. P. (2000), "Review: Web Surveys: A Review of Issues and Approaches*," *Public Opinion Quarterly*, 64, 464–494. https://doi.org/10.1086/318641.

Couper, M. P. (2007), "Issues of Representation in eHealth Research (with a Focus on Web Surveys)," *American Journal of Preventive Medicine*, Critical Issues in eHealth Research, 32, S83–S89. https://doi.org/10.1016/j.amepre.2007.01.017.

Data Reportal (2023), "Digital around the world."

Dutwin, D., and Buskirk, T. D. (2017), "Apples to Oranges or Gala versus Golden Delicious?: Comparing Data Quality of Nonprobability Internet Samples to Low Response Rate Probability Samples," *Public Opinion Quarterly*, 81, 213–239. https://doi.org/10.1093/poq/nfw061.

Fricker, R. D. (2017), "Sampling Methods for Online Surveys," in *The SAGE Handbook of Online Research Methods*, 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications Ltd, pp. 162–183. https://doi.org/10.4135/9781473957992.n10.

Glazer, J. V., MacDonnell, K., Frederick, C., Ingersoll, K., and Ritterband, L. M. (2021), "Liar! Liar! Identifying eligibility fraud by applicants in digital health research," *Internet Interventions*, 25, 100401. https://doi.org/10.1016/j.invent.2021.100401.

Goel, S., Obeng, A., and Rothschild, D. (2016), "Non-Representative Surveys: Fast, Cheap, and Mostly Accurate," *Working Paper*.

Keeter, S., Kennedy, C., Dimock, M., Best, J., and Craighill, P. (2006), "Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey," *Public Opinion Quarterly*, 70, 759–779. https://doi.org/10.1093/poq/nfl035.

Kennedy, C., Mercer, A., Keeter, S., Hatley, N., McGeeney, K., and Gimenez, A. (2016), "Evaluating Online Nonprobability Surveys."

Kohut, A., Keeter, S., Doherty, C., Dimock, M., and Christian, L. (2012), "Assessing the Representativeness of Public Opinion Surveys," *Pew Research Centre. For the People and the Press*.

Lee, S., Torok, M., Shand, F., Chen, N., McGillivray, L., Burnett, A., Larsen, M. E., and Mok, K. (2020), "Performance, Cost-Effectiveness, and Representativeness of

Facebook Recruitment to Suicide Prevention Research: Online Survey Study," *JMIR mental health*, 7, e18762. https://doi.org/10.2196/18762.

Leeper, T. J. (2019), "Where Have the Respondents Gone? Perhaps We Ate Them All," *Public Opinion Quarterly*, 83, 280–288. https://doi.org/10.1093/poq/nfz010.

Levy, P. S., and Lemeshow, S. (2013), *Sampling of Populations: Methods and Applications*, John Wiley & Sons.

Newson, J. J., Pastukh, V., and Thiagarajan, T. C. (2022), "Assessment of Population Well-being With the Mental Health Quotient: Validation Study," *JMIR Mental Health*, 9, e34105. https://doi.org/10.2196/34105.

Newson, J. J., Sukhoi, O., and Thiagarajan, T. C. (2024), "MHQ: constructing an aggregate metric of population mental wellbeing," *Population Health Metrics*, 22, 16. https://doi.org/10.1186/s12963-024-00336-y.

Newson, J. J., and Thiagarajan, T. C. (2020), "Assessment of Population Well-Being With the Mental Health Quotient (MHQ): Development and Usability Study," *JMIR Mental Health*, 7, e17935. https://doi.org/10.2196/17935.

Pedersen, E. R., and Kurz, J. (2016), "Using Facebook for Health-related Research Study Recruitment and Program Delivery," *Current opinion in psychology*, 9, 38–43. https://doi.org/10.1016/j.copsyc.2015.09.011.

Sanchez, C., Grzenda, A., Varias, A., Widge, A. S., Carpenter, L. L., McDonald, W. M., Nemeroff, C. B., Kalin, N. H., Martin, G., Tohen, M., Filippou-Frye, M., Ramsey, D., Linos, E., Mangurian, C., and Rodriguez, C. I. (2020), "Social media recruitment for mental health research: A systematic review," *Comprehensive Psychiatry*, 103, 152197. https://doi.org/10.1016/j.comppsych.2020.152197.

Sapien Labs (2021), *Mental State of the World 2020*.

Schneider, D., and Harknett, K. (2022), "What's to Like? Facebook as a Tool for Survey Data Collection," *Sociological Methods & Research*, 51, 108–140. https://doi.org/10.1177/0049124119882477.

Thornton, L., Batterham, P. J., Fassnacht, D. B., Kay-Lambkin, F., Calear, A. L., and Hunt, S. (2016), "Recruiting for health, medical or psychosocial research using Facebook: Systematic review," *Internet Interventions*, 4, 72–81. https://doi.org/10.1016/j.invent.2016.02.001.

Twenge, J. M., Cooper, A. B., Joiner, T. E., Duffy, M. E., and Binau, S. G. (2019), "Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005–2017.," *Journal of Abnormal Psychology*, 128, 185–199. https://doi.org/10.1037/abn0000410.

United Nations (2022), "World Population Prospects 2022."

US Census Bureau (2023a), "American Community Survey (ACS)."

US Census Bureau (2023b), "American Community Survey Data."

US Census Bureau (2024), "Household Pulse Survey: Measuring Emergent Social and Economic Matters Facing U.S. Households."

US Government Accountability Office (2002), "The American Community Survey: Accuracy and Timeliness Issues."

Wang, J., Calderon, G., Hager, E. R., Edwards, L. V., Berry, A. A., Liu, Y., Dinh, J., Summers, A. C., Connor, K. A., Collins, M. E., Prichett, L., Marshall, B. R., and Johnson, S. B. (2023), "Identifying and preventing fraudulent responses in online public health surveys: Lessons learned during the COVID-19 pandemic," *PLOS global public health*, 3, e0001452. https://doi.org/10.1371/journal.pgph.0001452.

Whitaker, C., Stevelink, S., and Fear, N. (2017), "The Use of Facebook in Recruiting Participants for Health Research Purposes: A Systematic Review," *Journal of Medical Internet Research*, 19, e7071. https://doi.org/10.2196/jmir.7071.