Working Paper 8:

Assessing the impact of video interviewing on measurement quality: Evidence from an experimental study on mode effects

Marc Asensio[1], Matt Brown[2], Richard Silverwood[2], Tim Hanson[3] & Gabriele Durrant[4]

[1]University of Lausanne; [2]University College London; [3]City St. George's, University of London; [4]University of Southampton

November 2025

Survey Futures is an Economic and Social Research Council (ESRC)-funded initiative (grant grant ES/X014150/1) aimed at bringing about a step change in survey research to ensure that high quality social survey research can continue in the UK. The initiative brings together social survey researchers, methodologists, commissioners and other stakeholders from across academia, government, private and not-for-profit sectors. Activities include an extensive programme of research, a training and capacity-building (TCB) stream, and dissemination and promotion of good practice. The research programme aims to assess the quality implications of the most important design choices relevant to future UK surveys, with a focus on inclusivity and representativeness, while the TCB stream aims to provide understanding of capacity and skills needs in the survey sector (both interviewers and research professionals), to identify promising ways to improve both, and to take steps towards making those improvements. Survey Futures is directed by Professor Peter Lynn, University of Essex, and is a collaboration of twelve organisations, benefitting from additional support from the Office for National Statistics and the ESRC National Centre for Research Methods. Further information can be found at www.surveyfutures.net.

Data collection was designed and managed by the UCL Centre for Longitudinal Studies and supported by the Economic and Social Research Council (ESRC) (Grant number ES/W013142/1). Fieldwork was conducted by Ipsos.

Prior to citing this paper, please check whether a final version has been published in a journal. If so, please cite that version. In the meanwhile, the suggested form of citation for this working paper is:

Asensio M., Brown M., Silverwood R., Hanson T. & Durrant G. (*2025*) ‘Assessing the impact of video interviewing on measurement quality: Evidence from an experimental study on mode effects’, *Survey Futures Working Paper* no. 8. Colchester, UK: University of Essex. Available at https://surveyfutures.net/working-papers/.

Table of Contents

Abstract

Amid declining response rates and rising survey costs, identifying reliable and cost-effective data collection methods is crucial. During the COVID-19 pandemic, major social surveys began to adopt video interviewing as an alternative to in-person data collection. However, its impact on measurement quality remains underexplored. This study addresses this evident gap by comparing mode effects between video, web, and in-person survey modes, using data from the first large-scale experiment where 1,692 respondents were randomly assigned to one of these modes. We focus on two key measurement quality indicators: item non-response and response distribution.

Our results show that video interviews yielded slightly lower item non-response levels than in-person interviews, albeit these differences are almost negligible. While measurement differences in responses to survey questions between the interviewer-administered modes were minimal, significant differences were observed between video and web responses, particularly on sensitive items like mental wellbeing and financial difficulties. Our findings suggest that video interviews are comparable to in-person surveys, but they may also suffer from the usual social desirability biases associated to the interviewer presence. Overall, these results are promising and suggest that video interviewing could serve as a valuable complement or alternative to traditional face-to-face surveys.

1. **Introduction**

The survey landscape has significantly evolved during the last decades, a rapid change mainly driven by technological advancements. The shift from interviewer-administered to self-administered web surveys, driven by the widespread adoption of the Internet, has been one of the most impactful developments. More recently, the proliferation of devices with integrated cameras and the rise of online video software—along with an increasing reliance on video technology for social and business communication—have made conducting video interviews[1] an accessible, affordable and feasible mode of survey data collection (Anderson 2008; Jeannis et al. 2013; Conrad et al. 2023; Endres et al. 2023).

Collecting survey data through video interviews may have several advantages. Anderson (2008) argues that, compared to self-administered modes of data collection, video interviews build on the presence of an interviewer, albeit remotely, to provide stronger rapport and higher survey satisfaction and, in parallel, persuade demographic groups typically underrepresented in surveys to participate. Moreover, offering video interviews as an additional (or alternative) survey mode could potentially increase response rates and decrease non-response biases (Jeannis et al. 2013; Durrant et al. 2025). Another important line of justification concerns the cost effectiveness of the method, as video interviews eliminate travel expenses and time associated with in-person interviews, allowing interviewers to conduct data collection without the financial costs of physically visiting respondents (Sullivan 2012; Janghorban et al. 2014).
These advantages have led researchers to suggest that video interviewing could be one promising solution to the many persistent challenges associated with in-person data collection, predominantly increasing survey costs and declining response rates (Endres et al. 2023; Schober et al. 2023; Centeno et al. 2024; Durrant et al. 2025) and could offer a cost-effective approach which retains the advantages (but also some limitations) of in-person interviewing (Sun et al. 2021; West et al. 2022).

The use of video interviewing for collecting data in a large-scale survey context was rare prior to the COVID-19 pandemic (Hanson et al. 2025). The restrictions following the lockdown prompted a rapid shift to remote data collection strategies and survey organisations relying on in-person data collection had to consider alternatives to their usual approaches. For example,

---

[1] When referring to ‘video interview’ we specifically refer to live video mediated interview as opposed to pre-recorded video interview, a self-administered alternative to the former.

in its tenth round, the European Social Survey introduced, for the first time, video interviewing as a complementary mode to in-person interviewing (Hanson et al. 2025). In the UK context, several major surveys adopted video interviewing as part of their data collection strategy, either as part of a mixed-mode approach (e.g. Next Steps, Children of the 2020s) or by offering video interview as the only method (e.g. 1970 British Cohort Study, 1958 National Child Development Study), at least for a period, when restrictions did not allow in-person contact (see Durrant et al. 2025 for more details).

The sudden shift in data collection methods left little room for testing how video interviewing might impact measurement and other data quality indicators, and as a result, this remains largely unknown (Endres et al., 2023). While emerging studies have provided some insights, many lack either sufficient external validity to generalise results or the necessary experimental design to disentangle the effect of self-selection into mode. For example, if a particular subpopulation is more likely to participate by video, then differences in measurement could be falsely attributed to the mode. Understanding whether and how the video interview mode influences measurement data quality is crucial to assess its potential as a mode for conducting surveys going forward.

To address this gap, this paper presents evidence from an experimental study in which respondents (N = 1,692) were randomly assigned to complete the same questionnaire via web, in-person interview, or video interview. We analyse mode effects by assessing differences in item non-response and response distribution across the three modes. Before presenting our research methods and results, in the following section we review the existing literature on video interview mode effects.

2. **Background**

Literature on the use of video interviews to collect large-scale survey data and its implications for data quality are still in its infancy. Despite this, survey practitioners are rapidly adapting. Some studies have already examined the effect of this mode on survey participation (Durrant et al. 2025; Kocar et al. 2025), while others have explored its impact on non-response biases and survey representativeness (West et al. 2022; Schober et al. 2023). In this section, we focus on the current understanding of video interview mode effects on measurement quality and review the few studies that exist in this area.

Some non-experimental studies have compared results between video and other modes based on various measures. In the Medical Expenditure Panel Survey, comparisons between video, in-person and telephone interview modes revealed that respondents participating via video were more likely to report using available records for completing the survey (e.g. payment records, prescription records, etc.) (Kelley et al. 2022). Hanson et al. (2025) compared item non-response rates, interview length and patterns of non-differentiation in interviews conducted by video and in-person across six countries participating in the European Social Survey and found only minor differences. Zavala-Rojas et al. (2023) investigated measurement invariance for two key European Social Survey (ESS) questionnaire measures: social trust and attitudes toward immigrants. They found no mode effects and concluded that data collected by video and in-person could be aggregated for analysis. Likewise, using data from the Italian Labour Force Survey, Rossetti et al. (2024) found no differences in employment status reporting between in-person and video interviews.

Few other studies have reported evidence based on experimental data. In a small-scale laboratory experiment conducted by Sun et al. (2021) where respondents were randomly assigned to either an in-person or a video interview, it was found that neither the disclosure of sensitive items nor the level of item non-response was significantly different between modes. Conrad et al. (2023) also assessed measurement quality in an experiment in which participants from a non-probability opt-in sample were randomly allocated to video interview or two self-administered modes—web survey and pre-recorded video interview. They found lower levels of non-differentiation in video interview responses, compared to responses from web and pre-recorded video respondents. However, respondents in the video interview mode were less likely to disclose sensitive information and exhibited higher item non-response rates for sensitive items. Despite this, they reported higher survey satisfaction. These findings align with expectations for in-person interviews, where past research has similarly highlighted differences between interviewer-administered and self-administered modes. As such, the author suggested that the key mechanism behind these differences was the absence or presence of the interviewer. However, this could not be empirically tested with the data available.

In a similar study, Endres et al. (2023) conducted an experiment in which they compared survey mode effects across video interviews, in-person interviews, and web surveys using data from a community research pool in a controlled lab setting. Their findings showed that video and in-person interviews were closely aligned on key measurement quality indicators including non-differentiation patterns, item non-response, and elaboration in open-ended responses. This led the authors to conclude that data collected by video more closely resembles data collected in-person than data collected in web-based self-administered surveys.

The current evidence points consistently to a high degree of similarity between data collected via video and data collected in-person. The cited studies demonstrate that video interviews share many of the advantages of in-person interviews, for example, higher survey satisfaction and reduced non-differentiation but also some of the drawbacks such as less disclosure of sensitive information due to social desirability bias. However, the existing evidence is highly limited and further experimental evidence is needed to deepen our understanding of the impact of video interviewing on survey responses (see discussion in Endres et al. 2023).

A particular aspect of the interview process that has been under-investigated in the video context is the use of self-completion modules. The aim of self-completion modules is to give participants greater privacy when responding to sensitive questions and to minimise the risk of social desirability bias that may occur if such questions are asked directly by an interviewer (Couper & Stinson 1999). However, there is evidence suggesting that the presence of an interviewer whilst participants are self-completing can still influence responses (Burkill et al., 2016). Given that the use of self-completion modules is standard practice in interview-administered modes, it is also important to understand better how it works within the video context and how it compares to other modes.

Building on existing research, this paper contributes to the ongoing debate by presenting the findings from the first large-scale experimental study comparing measurement quality in video interviews with in-person and web modes.

3. **Methods**

3.1. **Data**

We use data from a survey mode experiment conducted in England by the Centre for Longitudinal Studies (CLS). CLS runs a series of longitudinal cohort studies including Next Steps, which is following the lives of around 16,000 people born in England in 1989-1990 and the Millennium Cohort Study, which is following around 19,000 people born in the United Kingdom in 2000-2002 (Conelly et al. 2014; Joshi et al. 2016; Wu et al. 2024). Both studies used a web-first mixed mode data collection strategy in the most recently completed waves – the Next Steps Age 32 Survey and the Millennium Cohort Study Age 23 Survey. The mode experiment was conducted to assess the impact of survey mode (web, video or in-person interview) on some of the key measures included in these surveys. It was conducted with a newly recruited, independent sample, rather than participants from the two studies.

The study targeted residents of England aged 20 to 40 to ensure comparability with the Next Steps and Millennium Cohort Study participants at the time of the recently completed waves. Participants were recruited on a free-find basis through in-home and in-street approaches in 200 sample points which were designed to reflect the regional distribution of the population of England. To maintain a broadly representative sample in terms of gender, age, and region, soft quotas were applied. In total, 1,800 participants were recruited (Appendix 1). During pre-recruitment, individuals provided contact details and consent to participate in the study. Individuals were given a participant information sheet describing the study, a privacy notice and a thank you note with their interviewer contact details. Importantly, at the time of the pre-recruitment respondents were not made aware of the mode they had been assigned to. Participants were then randomly allocated, using a rotation plan, to one of three experimental mode groups (web, in-person, or video) based on the order in which they were recruited.
Consent to participate was re-confirmed two weeks before the start of the fieldwork period. Fieldwork took place from January 17th to February 9th, 2024, and respondents were offered a conditional £20 voucher for their participation. A total of 1,692 respondents took part in the survey. Appendix 2 presents the demographic characteristics and distribution of the participants.

At the outset of fieldwork, respondents assigned to the web mode received a survey link via email and completed the questionnaire independently. Those allocated to in-person and video modes were contacted by interviewers at the beginning of fieldwork to arrange appointments for the interview. The same team of interviewers conducted both in-person and video interviews; there were no separate interviewer groups dedicated to each mode. In the in-person mode, interviews were conducted at respondents' homes with the interviewer recording responses on a tablet. Respondents could not see the tablet screen, and showcards were provided when necessary. In the video mode, interviews were conducted via Microsoft Teams. Interviewers shared their screen with the respondents and read the questions aloud.

The survey contained seven different modules: (1) household relationships, (2) housing, (3) activities and employment, (4) finances, (5) health, (6) identity and (7) the self-completion module. The self-completion module included questions on financial literacy, life satisfaction, mental well-being, drinking and smoking behaviour, gender identity, adverse childhood circumstances and a cognitive assessment. For the web mode, which is entirely self-completed, this module was simply a continuation of the survey. In the in-person mode, the interviewer handed their device to the respondent to complete this last module. Respondents completing video interviews were provided with a link to complete a separate online survey during the interview. The interviewers remained on the Teams call but stopped sharing their screen. To discourage the use of the "Don't know" and "Refusal" options, these were not explicitly offered in any of the three modes. In the web mode, respondents had to click "Next" without answering for these options to appear. Similarly, in the video mode, these options were only shown on the interviewer's screen if no response was given and "Next" was selected. In the in-person mode, the options were not read aloud but were recorded if spontaneously mentioned by the respondent.

3.2. **Analytical approach**

In this study, we assess mode effects on two key measurement quality indicators: item non-response and differences in response distribution. Since significant differences were observed between modes for several demographic variables, specifically gender, education, and social grade (Appendix 2), we employ a regression-based approach to control for these and other socio-demographic variables (detailed later in this section). We analyse *all* questionnaire items, excluding those related to occupational coding and cognitive test experiments (which will be analysed in detail in separate publications) *and* socio-demographic items which are used as controls. In total, our analysis covers 25 items (see Appendix 3 for details).

To evaluate **item non-response** differences, we first created a binary variable for each item, indicating whether a respondent provided an answer or not. Non-response was defined as respondents reporting that they did not know the answer or did not wish to answer. Whilst some "I do not know" answers might be genuine, differences in reporting patterns might be, according to de Leeuw et al. (2016), attributed to higher levels of measurement error in a certain survey mode. Questions that were answered by all respondents across all modes (2 variables) or by all respondents in the video mode only (2 variables) were excluded from the item non-response analysis. Additionally, we excluded 12 more variables with a low missing rate, i.e.

with 10 or fewer missing observations, as they fell below the commonly accepted threshold of events per predictor variable (Vittinghoff & McCulloch 2007). Hence, in total, 16 variables were excluded from this item non-response analysis (see Appendix 3 for a detailed list).

For the remaining 9 variables, we fitted logistic regression models, modelling item non-response as a function of survey mode (video interview (reference category), web or in-person interview) and controlling for the socio-demographic variables listed in Appendix 2. These included: sex (male, female), age, ethnicity (white, non-white), highest educational attainment (GCSE or lower, further education, higher education), employment status (in paid work, other), social grade (AB, C1, C2, DE), living as a couple (yes, no), having children (yes, no), housing tenure (own, rent, rent-free & other) and region (Greater London, Nort West, North East, Yorkshire & Humber, West Midlands, East Midlands, South East, East of England, South West). To streamline the results and improve readability, we present in the results section only the coefficient estimates (with 95% confidence intervals) for the mode variable—one coefficient for video interview vs. web and another for video interview vs. in-person. Readers interested in the full model results for all 9 variables, can refer to SOM1. We present the variables which were part of the self-completion module in a separate figure.

To provide a complete picture of item non-response across modes, we also computed three count variables indicating the number of occasions where a respondent failed to provide an answer to (1) the items from first six modules (*main questionnaire* modules from now on); (2) the items from the self-completion module and (3) the total summing the previous two. We modelled the number of item non-responses again as a function of mode and the control variables, as specified in the above paragraph, using a Poisson model for each count variable (full model specifications and results in SOM2).

To analyse differences in **response distributions** across modes, we regressed each of the 25 survey variable on mode and the set of socio-demographic controls, as specified in the previous analysis. We used three different types of models[2]: Poisson for count variables, logistic for

[2] Planned analyses involved different models being applied depending on the variable's nature. However, the high number of ordinal variables complicated this approach. Some required a binary transformation due to the marginal number of cases in some categories, others violated the proportional odds assumption required for ordinal logistic models (and as such, a generalised ordinal logistic model or multinomial logistic model would be more appropriate) and the rest, did not. In order to reduce the total number of different models used and to simplify presentation, we used binary transformations for all ordinal variables

binary variables (including transformed ordinal variables) and finally, a linear model for one variable that was measured on a 10-point scale. Appendix 3 summarises the binary transformations and the models used to regress each variable. To maximise clarity, we present coefficient estimates (with 95% confidence intervals) for the mode comparison only, while complete model specifications with results of all coefficients are provided in SOM3. Again, we present the results distinguishing between items from the main questionnaire and those from the self-completion module.

Due to the minimal number of missing cases in our control variables (see Appendix 2), we used a complete-case variable-specific approach for our regression models.

4. Results

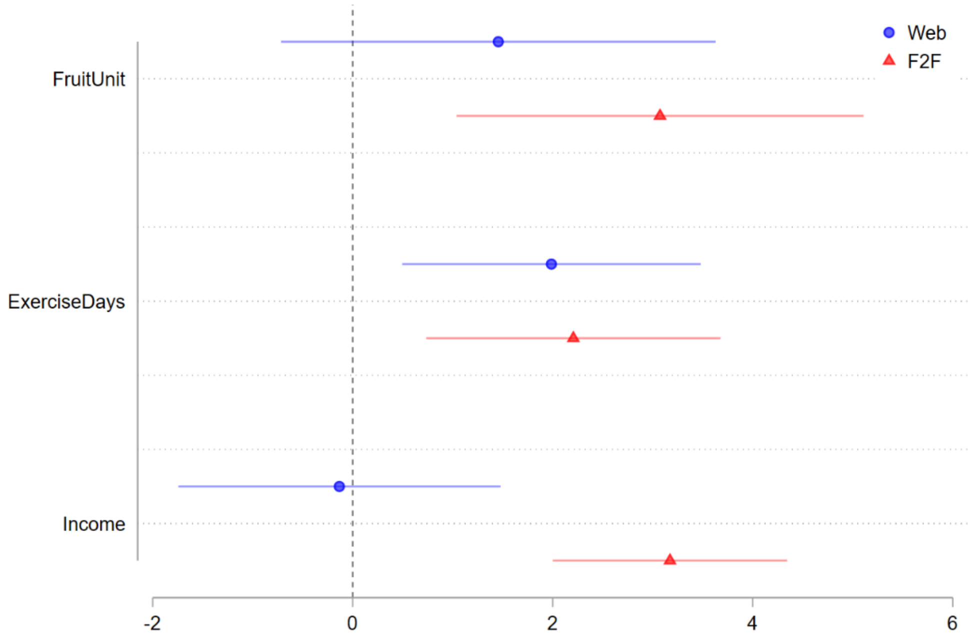
Item non-response analysis

Item non-response was rare across all three modes for the main questionnaire modules. The average number of total missing items observed (non-adjusted) was 1.74 for in-person respondents, 1.40 for video respondents, and 1.61 for web respondents out of a total of 25 items. As a result, many variables did not have enough missing observations to conduct a meaningful item non-response analysis.

Only three variables in the first six questionnaire modules surpassed the events per predictor threshold (Figure 1): number of fruits eaten per day, number of days doing exercise per week and income. Respondents assigned to the in-person mode were significantly more likely not to answer these items compared to video interview respondents. However, no such difference was observed when comparing web and video interviews. The only exception was for the exercise frequency variable, where web respondents were more likely not to answer.

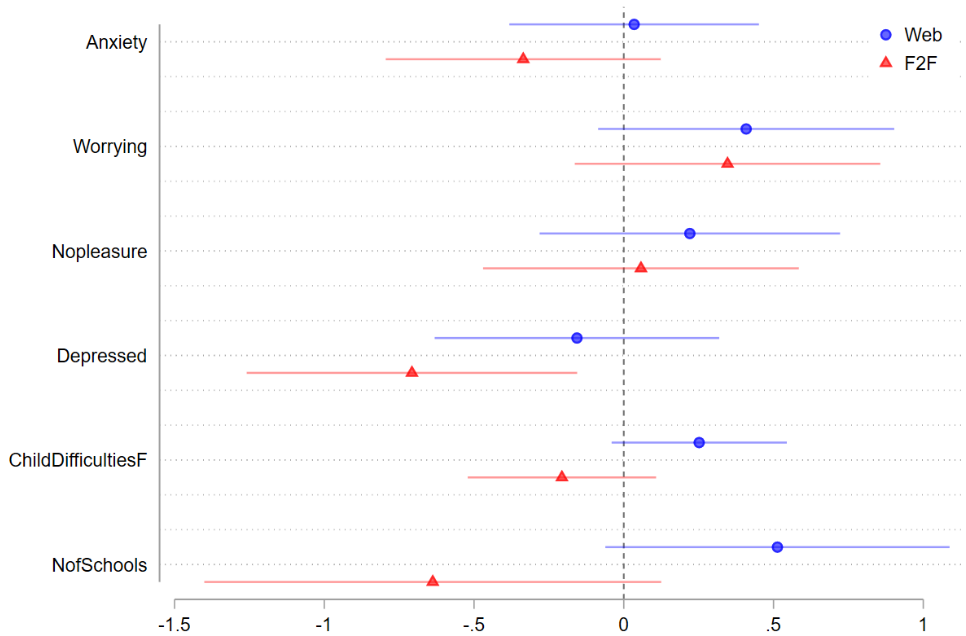
The self-completion module contained more sensitive items and as such, item non-response was more common for these variables (Figure 2). However, only one significant difference was observed across both mode comparisons: video interview respondents were less likely to answer the depression item than those interviewed in-person.

Figure 1. Mode coefficient estimates from the logistic regression models for item non-response for variables in the main questionnaire modules



Note: 95% confidence intervals. Video mode as reference category.

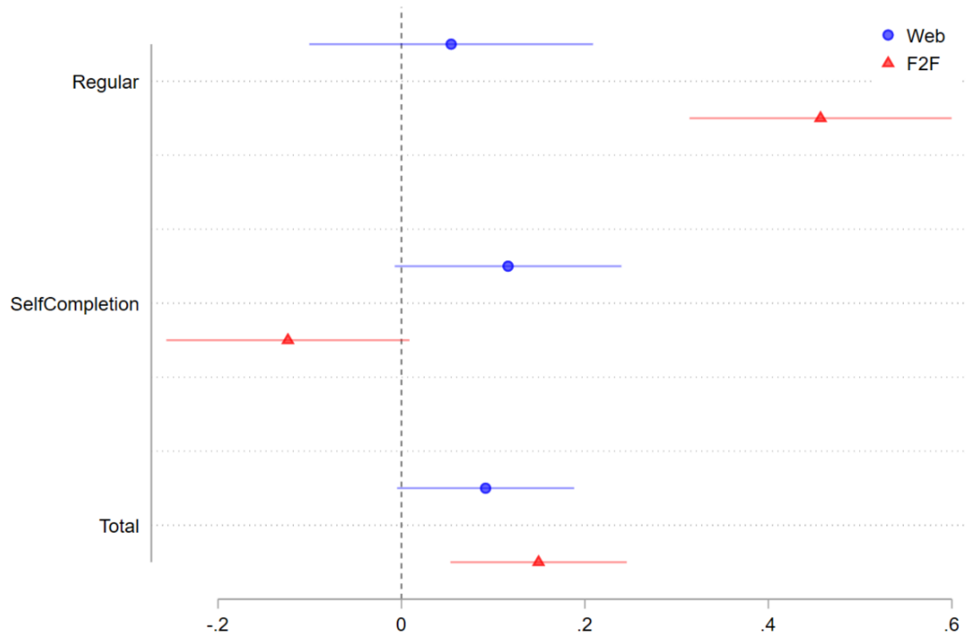
Figure 2. Mode coefficient estimates from the logistic regression models for item non-response for the variables in the self-completion module



Note: 95% confidence intervals. Video mode as reference category.

For a comprehensive overview of item non-response throughout the questionnaire, we also model the derived count variable, that indicates the total number of missing items for a particular respondent, using Poisson regression. The results are presented in Figure 3. In the main questionnaire modules, respondents assigned to the in-person interview had, on average, 63% more missing answers than video interview respondents ($\beta = 0.49$; IRR = 1.63 $p < .001$, which is in line with the results from the item-specific analysis from the main questionnaire modules variables. In addition, for the self-completion module, there was some indication of differences in item non-response across modes, but these did not reach conventional levels of statistical significance. When combining all modules, in-person respondents had, on average, 16% more missing items than those assigned to the video group.

Figure 3. Mode coefficient estimates from the Poisson regression models for number of missing items for the main questionnaire modules, self-completion module and the total.



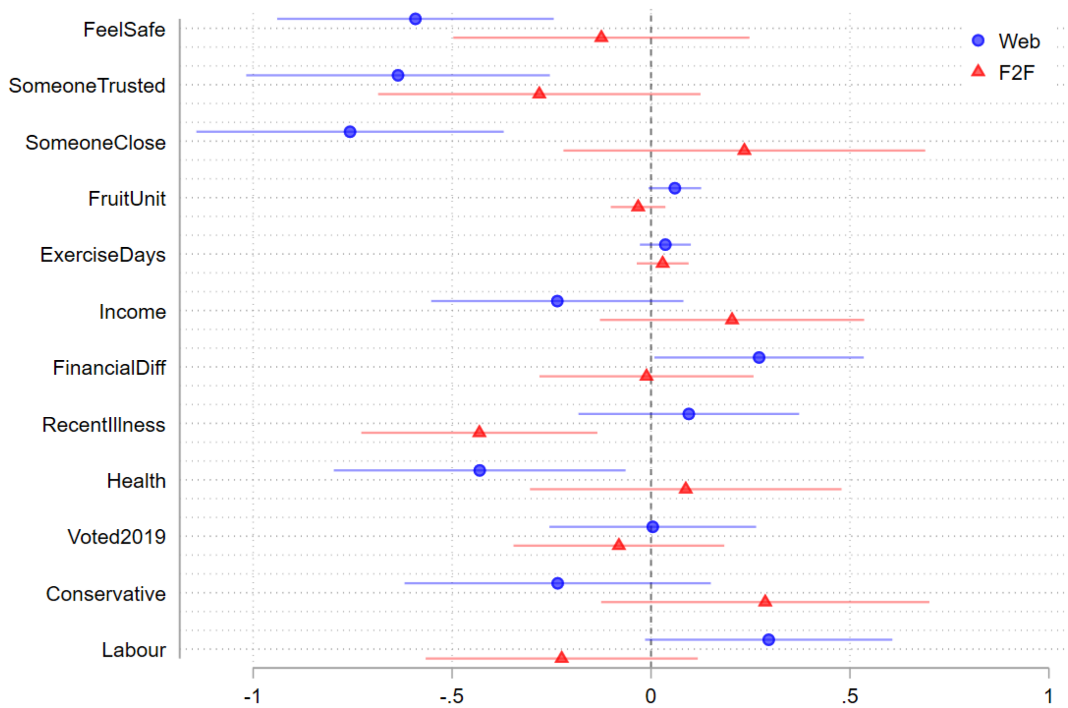
Note: 95% confidence intervals. Video mode as reference category.

Response distribution analysis

For the main questionnaire modules, we observe no significant measurement differences between the interviewer-administered modes (Figure 4), except for the illness variable, where in-person respondents were less likely to report having had a recent illness. However, more differences emerge when comparing web and video interviews: web respondents were more likely to report feeling less safe and were also more prone to indicate that they lacked someone to trust or someone close. Additionally, they were more likely to report financial difficulties and lower levels of health.

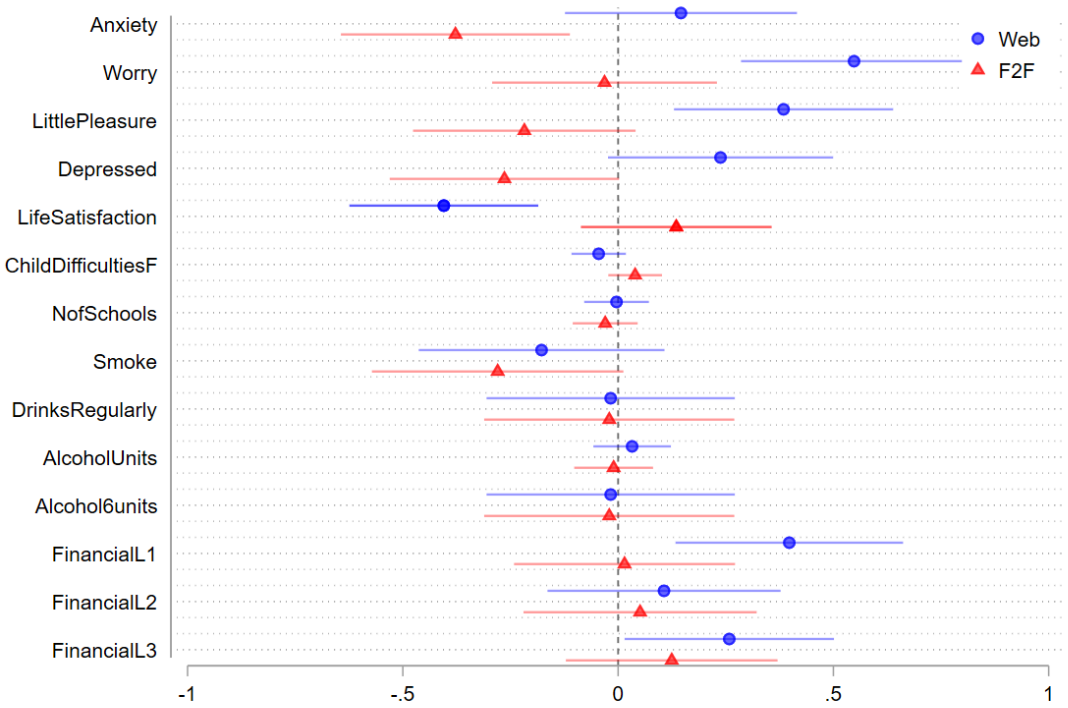
A similar pattern is observed for the self-completion module (Figure 5). Compared to video interview respondents, web respondents were more likely to report higher levels of mental distress, including feeling more anxious, more worried, finding little pleasure in doing things and lower levels of life satisfaction. In addition, they were more likely to answer correctly on two out of the three financial literacy questions. In contrast, for the comparison between the video and in-person interviews, only one significant difference was observed: in-person respondents reported lower levels of anxiety.

Figure 4. Mode coefficient estimates from the regression models for the outcome variables for the main questionnaire modules



Note: 95% confidence intervals. Video mode as reference category.

Figure 5. Mode coefficient estimates from the regression models for the outcome variables for the self-completion module



Note: 95% confidence intervals. Video mode as reference category.

5. **Discussion**

During the COVID pandemic there was a surge in the use of video-interviewing to collect large-scale survey data. In many cases the mode was rapidly introduced with little time to fully evaluate the impact that this new mode might have on measurement quality. Now, with a growing need for cost-effective alternatives to in-person surveys, evaluating video interviewing as a mode of data collection in the post-pandemic context is of paramount importance. This study examines video interview mode effects using data from an experimental survey in which respondents were randomly assigned to complete the same questionnaire via either a self-administered online mode (web) or an interviewer-administered mode (in-person or video interview). We compare measurement quality across modes using two indicators: item non-response and response distribution.

5.1. **Main results**

Our results showed higher levels of item non-response for in-person interviews, compared to video interviews, contradicting recent literature on the topic where either no differences were found (Sun et al. 2021; Endres et al. 2023; Hanson et al. 2025) or higher levels were found in the video mode (Conrad et al. 2023). However, it is important to acknowledge the relatively low level of item non-response throughout the survey. For only one variable related to financial difficulties during childhood was item non-response higher than 6% (more specifically, it was 20%). We assume this was most likely due to a respondents' genuine lack of knowledge about the measure rather than satisficing behaviour (de Leeuw et al. 2016).

Practically speaking, item non-response was only marginally higher in the in-person interviews. Interestingly, the treatment of the "Don't know" and "Refusal" options followed the same principle across all modes: these options were not explicitly presented to discourage their use. The video interview mode combined features of both the web and in-person modes. While interviewers read aloud all response options, these were simultaneously displayed on the respondent's screen. Despite this hybrid setup, our results indicate that patterns of item non-response in the video mode were more closely aligned with those observed in the self-administered web mode.

Regarding response distribution differences, we found very few measurement differences between the interviewer-administered modes, in line with recent literature on the topic (Conrad et al. 2023; Endres et al. 2023; Zavala-Rojas et al. 2023; Rossetti et al. 2024). Out of the 25 items analysed in this study, we found significant differences between interviewer-administered modes only for two. On the other hand, we found many more differences between the self-administered web mode and video, particularly on sensitive items such as mental wellbeing, health and financial difficulties. These findings suggest that respondents assigned to both interviewer-administered modes were affected by social desirability bias and were less likely to report themselves in a negative light, even when these questions were part of the self-administered module. Our results support the conclusion that the key mechanism for this dynamic is the presence of the interviewer, despite the interaction being mediated by an online video platform (Conrad et al. 2023).

Importantly, we also observed differences between the video and the web groups for the financial literacy test questions, with web respondents being more likely to get the right answer. A possible explanation for this is that online respondents may have looked up answers online or used a calculator. In-person and video interviews contained a section conducted by self-completion. This approach is taken to increase privacy and mitigate the potential for social desirability bias. However, our findings suggest that the presence of an interviewer during completion of this section, either virtually or in-person, can still make participants more likely to provide socially desirable responses. Altogether, our results are mostly consistent with the literature on video interview as a mode for survey administration and the large body of literature on interviewer effects on social desirability.

5.2. **Limitations and future research**

Our study is not without its limitations. First, we used a non-probability-based sample of adults aged 20 to 40, which limits generalisation to the whole population. Other age groups may face challenges related to technology unfamiliarity or hearing difficulties, potentially leading to a different survey experience. Nonetheless, the randomization of the sample and our analytical approach accounting for differences in socio-demographic groups should provide a strong foundation for extrapolating these findings to similar contexts (e.g., Mullinix et al. 2015).
Second, our findings represent only one piece of the larger puzzle on how mode effects influence measurement quality. To gain a more comprehensive understanding, other indicators (e.g., non-differentiation patterns, elaboration of open-ended answers or scales validity *and* reliability) could also be explored to determine whether interviewer-administered modes differ in aspects beyond those we examined.

Third, it could be argued that items with higher levels of non-response – especially when differences between modes were observed – may introduce selection effects that could bias response distributions. In our analyses, for variables where item non-response differed significantly across modes, we did not observe corresponding mode differences in the substantive distributions. Furthermore, any possible selection effect was likely mitigated through the adjustment for control variables we followed in our analyses, so we believe this only constitutes a minor limitation.

Future research should test the robustness of our results and, ideally, expand attention to broader populations while accounting for the limitations and concerns described in this section. We believe there are still significant contributions to be made to the topic, in particular by exploring measurement quality in video interviews featuring different designs.

5.3. **Conclusions and practical recommendations**

Our study has several implications for the use of video- interviews as a mode of data collection. First, our results suggest that data collected by video is broadly comparable to data collected in-person. Video respondents appeared slightly more willing to respond to certain survey questions than those participating in-person, and amongst those answering we find very limited evidence of mode effects on responses between these interviewer-administered modes. This is a promising finding which suggests that video interviews could serve as an effective alternative (or complement) to in-person interviews without sacrificing data quality. Yet, survey practitioners should also consider the recruitment aspect (Anderson 2008). Recent research has demonstrated that persuading respondents to participate in video interviews can be challenging (see Guggenheim et al. 2021; Conrad et al. 2023, Hupp et al. 2024; Durrant et al. 2025). Analysis of the tenth round of the European Social Survey (Hanson et al. 2025), where video uptake varied substantially across countries, suggests that video interviews might not be an alternative to in-person interviews for all situations and contexts, but rather for specific ones. In cases where video interviews are deemed an optimal survey strategy, our results are reassuring for in-person interview data comparability but less so for web surveys.

Second, video interviews, much as in-person interviews, suffer from social desirability bias. This negatively affects the quality of sensitive measures collected, especially when compared to web surveys. Interviewer administered surveys often use self-completion modules to ask sensitive questions so as to grant additional privacy and mitigate the risk of social desirability bias. However, our results also demonstrate that this method is not perfect and that the interviewer's presence, whether in-person or by video still appeared to influence respondent reporting. One practical advantage of video over in-person is that the video interviewer could turn-off their camera and make their presence less obvious for the respondent or even terminate the call to provide more privacy to the respondent. We believe video interviews, due to certain similarities with web surveys, can offer more tools for dealing with this issue. However, whether video interviews as a mode can effectively bridge this obstacle remains to be seen.

Third, our analysis of the financial literacy test revealed higher scores among web respondents, raising the possibility that some may have consulted external resources while completing the test. This finding has important implications for the administration of certain tasks or tests and the collection of complex elements in self-administered contexts. For some time, survey practitioners have struggled with implementing certain elements in web contexts with the same success they have achieved in in-person interviewer administered surveys, particularly as the costs of the latter increase (e.g., occupational coding, collecting consent for data linkage, cognitive testing). A possible solution to this problem could be the use of video interviews, where leveraging social presence and interviewer-respondent rapport might achieve results similar to in-person surveys at a lower cost. Yet, this solution remains contingent on video interview participation and representation, and its success in these aspects is currently uncertain.

Fourth, our findings may be attributed to the hybrid nature of the video mode used in our design, which combined elements of both web and in-person survey modes. While interviewers read the questions aloud, respondents were able to view the full web survey interface on their screens. In contrast, other studies have adopted a more in-person-like approach, where interviewers shared their screen only to display specific showcards rather than the entire questionnaire. These different configurations highlight the design flexibility of video interviews, allowing researchers to selectively incorporate features from other modes to enhance response quality or survey experience. Exploring other design alternatives is of paramount importance for the optimisation of the video mode.

Altogether, our findings suggest that the measurement quality across interviewer-administered modes is broadly comparable which points to promising opportunities for the wider adoption of video interviewing in survey research, particularly as both survey organisations and respondents become more accustomed to the mode. However, current challenges in implementing video interviews, including technical barriers, interviewer training and challenges with respondent recruitment must not be overlooked.

6. **References**

Anderson, A. H. (2008). Video-Mediated Interactions and Surveys in *Envisioning the Survey Interview of the Future*, eds. Frederick G. Conrad and Michael F. Schober, Chapter 5, pp. 95–118, New York: Wiley & Sons. Belli, R. F., I. Bilgen, and T. Al Baghal

Centeno, L., Arrue, J., Good, C., Edwards, B., & Steiger, D. (2024). Practical considerations for developing and implementing a video mode for survey data collection. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, *164*(1), 122-146.

Connelly, R., & Platt, L. (2014). Cohort profile: UK Millennium Cohort Study (MCS). *International Journal of Epidemiology, 43*(6), 1719–1725.

Conrad, F. G., Schober, M. F., Hupp, A. L., West, B. T., Larsen, K. M., Ong, A. R., & Wang, T. (2023). Video in survey interviews: Effects on data quality and respondent experience. *Methoden, daten, analysen*, *17*(2), 135.

Couper, M. P., & Stinson, L. L. (1999). Completion of self-administered questionnaires in a sex survey. *Journal of Sex Research*, *36*(4), 321-330.

de Leeuw, E. D., Hox, J. J., & Boevé, A. (2016). Handling do-not-know answers: exploring new approaches in online and mixed-mode surveys. *Social Science Computer Review*, *34*(1), 116-132.

Durrant, G., Kocar, S., Brown, M., Hanson, T., Sanchez, C., Wood, M., Taylor, K., Tsantani, M., Huskinson, T. (2025). Live video interviewing: evidence of opportunities and challenges across seven major UK Social Surveys. Working paper: https://surveyfutures.net/wp-content/uploads/2024/06/working-paper-01-livevideo-interviewing.pdf.

Endres, K., Hillygus, D. S., DeBell, M., & Iyengar, S. (2023). A randomized experiment evaluating survey mode effects for video interviewing. *Political Science Research and Methods*, *11*(1), 144-159.

Hanson, T., Briceno-Rosas, R., Ainsaar, M., Kartau, A.K.: Live Video Interviewing as a Complementary Mode to In-Person Interviews: Evidence from the European Social Survey. Working paper: https:// surveyfutures.net/wp-content/uploads/2024/07/working-paper-02-livevideo-interviewing.pdf. (2024)

Höhne, J. K., Ziller, C., & Lenzner, T. (2024). Investigating respondents' willingness to participate in video-based web surveys. *International Journal of Market Research*, *66*(1), 3-13.

Hupp, A., Guggenheim, L., & Howell, D. (2024). Investigating Participation in Live Video Interviews. In *79th Annual AAPOR Conference*. AAPOR.

Irani, E. (2019). The use of videoconferencing for qualitative interviewing: Opportunities, challenges, and considerations. *Clinical Nursing Research*, *28*(1), 3-8.

Janghorban, R., Roudsari, R. L., & Taghipour, A. (2014). Skype interviewing: The new generation of online synchronous interview in qualitative research. *International journal of qualitative studies on health and well-being*, *9*(1), 24152.

Jeannis, M., Terry, T., Heman-Ackah, R., & Price, M. (2013). Video interviewing: An exploration of the feasibility as a mode of survey application. *Survey Practice*, *6*(1).

Joshi, H., & Fitzsimons, E. (2016). The Millennium Cohort Study: The making of a multi-purpose resource for social science and policy. *Longitudinal and Life Course Studies, 7*(4), 409–430.

Mullinix, K. J., Leeper, T. J., Druckman, J. N., & Freese, J. (2015). The generalizability of survey experiments. *Journal of Experimental Political Science*, *2*(2), 109-138.

Phillips, B., Neiger, D., Slamowicz, S., Lester, G., Luddon, S., Farrell, E., Gerlach, K., & Carmo, P. (2023). Video-Assisted Live Interviewing in Comparison to Other Survey Methods in Australia [Conference presentation]. 10th Conference of the European Survey Research Association. Milan, Italy.

Rossetti, F., Bontempi, K., & Pietropaoli, S. (2024). A comparison between Video Mediated Interview (VMI) and CAPI in Labour Force Survey. *METRON*, 1-17.

Schober, M. F., Conrad, F. G., Hupp, A. L., Larsen, K. M., Ong, A. R., & West, B. T. (2020). Design considerations for live video survey interviews. *Survey Practice*, *13*(1).

Schober, M. F., Okon, S., Conrad, F. G., Hupp, A. L., Ong, A. R., & Larsen, K. M. (2023). Predictors of Willingness to Participate in Survey Interviews Conducted by Live Video. *Technology, Mind, and Behavior*, *4*(2: Summer 2023).

Sullivan, J. R. (2012). Skype: An appropriate method of data collection for qualitative interviews?. *The Hilltop Review*, *6*(1), 10.

Sun, H., Conrad, F. G., & Kreuter, F. (2021). The relationship between interviewer-respondent rapport and data quality. *Journal of Survey Statistics and Methodology*, *9*(3), 429-448.

Vittinghoff, E., & McCulloch, C. E. (2007). Relaxing the rule of ten events per variable in logistic and Cox regression. *American journal of epidemiology*, 165(6), 710-718.

West, B. T., Ong, A. R., Conrad, F. G., Schober, M. F., Larsen, K. M., & Hupp, A. L. (2022). Interviewer effects in live video and prerecorded video interviewing. Journal of Survey Statistics and Methodology, *10*(2), 317-336.

Wu, A. F., Henderson, M., Brown, M., Adali, T., Silverwood, R. J., Peycheva, D., et al. (2024). Cohort profile: Next Steps—the longitudinal study of people in England born in 1989–90. *International Journal of Epidemiology, 53*(6).

Zavala-Rojas, D., Keles, O.K., Schwarz, H., & Romanin, E. (2023). Measurement invariance of two concepts across face-to-face and video interview mode in ESS R10. 10th Conference of the European Survey Research Association. Milan, Italy.

7. **Appendices**

Appendix 1. Sample sizes across the sample points by region.

Region	**% aged 20-40 in England**	**Sampling Points**	**Sample Size**
North West	13.12%	26	229
North East	4.81%	10	82
Yorkshire & The Humber	8.77%	18	173
West Midlands	9.94%	20	188
East Midlands	8.56%	17	147
South East	15.22%	30	268
East of England	10.24%	20	186
South West	9.49%	19	161
Greater London	19.85%	40	367
Total	100%	200	1.800

Appendix 2. Distribution of socio-demographic variables

Variable		Total	Web	F2F	Video	p-value
	Sample Size	1,692 (100%)	582 (34.40%)	565 (33.39%)	545 (32.21%)	
Gender	Male	753 (44.50%)	294 (50.52%)	226 (40.0%)	233 (42.75%)	<0.01
	Female	919 (54.31%)	278 (47.77%)	337 (59.65%)	304 (55.78%)	
	Other	20 (1.14%)	10 (1.72%)	2 (0.36%)	8 (1.47%)	
Age	20-30	779 (46.04%)	296 (50.86%)	259 (45.84%)	224 (41.10%)	<0.01
	31-41	913 (53.96%)	286 (49.14%)	306 (54.16%)	321 (58.90%)	
Ethnic Group	White	1,549 (91.55%)	525 (90.21%)	523 (92.57%)	501 (91.93%)	
	Non-White	139 (8.22%)	54 (9.28%)	42 (7.43%)	43 (7.89%)	
	Missing	4 (0.24%)	3 (0.52%)	0 (0%)	1 (0.18%)	
Education	GCSE or lower	471 (27.84%)	154 (26.46%)	178 (31.50%)	139 (25.50%)	<0.01
	Further education	827 (48.88%)	291 (50.0%)	280 (49.56%)	256 (46.97%)	
	Higher education	394 (23.29%)	137 (23.54%)	107 (18.94%)	150 (27.52%)	
Employed	In paid work	1,212 (71.63%)	421 (72.34%)	390 (69.03%)	401 (73.58%)	
	Other	478 (28.25%)	160 (27.49%)	174 (30.80%)	144 (26.42%)	
	Missing	2 (0.12%)	1 (0.17%)	1 (0.18%)	0 (0%)	
Social Grade	AB	298 (17.61%)	99 (17.01%)	88 (15.58%)	111 (20.37%)	<0.01
	C1	631 (37.29%)	210 (36.08%)	205 (36.28%)	216 (39.63%)	
	C2	389 (22.99%)	159 (27.32%)	131 (23.19%)	99 (18.17%)	
	DE	374 (22.10%)	114 (19.59%)	141 (24.96%)	119 (21.83%)	
Living as couple	Yes	974 (57.57%)	324 (55.67%)	333 (58.94%)	317 (58.17%)	
	No	715 (42.26%)	257 (44.16%)	230 (40.71%)	228 (41.83%)	
	Missing	3 (0.18%)	1 (0.17%)	2 (0.35%)	0 (0%)	
Has any child	Yes	968 (57.21%)	320 (54.98%)	332 (55.76%)	316 (57.98%)	
	No	723 (42.73%)	262 (45.02%)	232 (41.06%)	229 (42.02%)	
	Missing	1 (0.06%)	0 (0%)	1 (0.18%)	0 (0%)	
Tenure	Own	620 (36.64%)	200 (34.36%)	215 (38.05%)	205 (37.61%)	
	Rent	711 (42.02%)	243 (41.75%)	228 (41.35%)	240 (44.04%)	

	Rent-free & other	349 (20.63%)	137 (23.54%)	112 (20.82%)	100 (18.35%)
	Missing	10 (0.59%)	2 (0.34%)	8 (1.42%)	0 (0%)
Region	North West	246 (14.54%)	82 (14.09%)	84 (14.87%)	90 (14.68%)
	North East	82 (4.85%)	30 (5.15%)	28 (4.96%)	24 (4.40%)
	Yorkshire & Humbersid	153 (9.04%)	52 (8.93%)	52 (9.20%)	49 (8.99%)
	West Midlands	185 (10.93%)	61 (10.48%)	62 (10.97%)	62 (11.38%)
	East Midlands	141 (8.33%)	49 (8.42%)	48 (8.50%)	44 (8.07%)
	South East	255 (15.07%)	87 (14.95%)	88 (15.58%)	80 (14.68%)
	East of England	157 (9.28%)	56 (9.62%)	50 (8.85%)	51 (9.36%)
	South West	150 (8.87%)	53 (9.11%)	49 (8.67%)	48 (8.81%)
	Greater London	323 (19.09%)	112 (19.24%)	104 (18.41%)	107 (19.63%)

Appendix 3. Overview of information for variables from main questionnaire and self-completion modules (variables of substantive interest): frequency distribution and information on item non-response modelling.

Main questionnaire modules

Variable and codename	Categories	Observations	Binary transformation	Model	Item non-response analysis
Feels safe, secure and happy "FeelSafe"	Very true Partly true Not at all true N = 1,689 Missing = 0.18%	1,438 (85.14%) 238 (14.09%) 13 (0.77%)	Very true vs. rest	Logistic	Excluded
Someone to trust "SomeoneTrusted"	Very true Partly true Not at all true N = 1,691 Missing = 0.06%	1,484 (87.76%) 189 (11.18%) 18 (1.06%)	Very true vs. rest	Logistic	Excluded
No one I feel close to "SomeoneClose"	Very true Partly true Not at all true N = 1,687 Missing = 0.30%	61 (3.62%) 122 (7.23%) 1,504 (89.15%)	Not at all true vs. rest	Logistic	Excluded
Fruits and vegetables eaten in a typical day "FruitUnit"	0 1 2 3 4 5 6 7 8 9 10 12 14 15 30 50 N = 1,661 Missing = 1.83%	62 (3.73%) 228 (13.73%) 377 (22.70%) 456 (27.45%) 210 (12.64%) 232 (13.97%) 39 (2.35%) 19 (1.14%) 12 (0.72%) 1 (0.06%) 14 (0.84%) 2 (0.12%) 1 (0.06%) 2 (0.12%) 5 (0.30%) 1 (0.06%)	NA	Poisson	Yes
Number of days per week do exercise for 30 minutes or more "ExerciseDays"	0 1 2 3 4 5 6	205 (12.40%) 132 (7.99%) 232 (14.04%) 290 (17.54%) 192 (11.62%) 327 (19.78%) 77 (4.66%)	NA	Poisson	Yes

	7 N = 1,653 Missing = 2.30%	198 (11.98%)			
Household income “Income”	1st bracket 2nd bracket 3rd bracket 4th bracket N = 1,618 Missing = 2.84%	215 (13.29%) 616 (38.07%) 491 (30.35%) 296 (18.29%)	1st and 2nd bracket vs the rest	Logistic	Yes
How managing financially these days “FinancialDiff”	Living comfortably Doing alright Just about getting by Finding it quite difficult Finding it very difficult N = 1,685 Missing = 0.41%	221 (13.12%) 772 (45.82%) 480 (28.49%) 148 (8.78%) 64 (3.80%)	Alright and comfortably vs the rest	Logistic	Excluded
General state of health “Health”	Excellent Very good Fair Poor N = 1,690 Missing = 0.12%	285 (16.86%) 629 (37.22%) 556 (32.90%) 47 (2.78%)	Very good and excellent vs. the rest	Logistic	Excluded
Any illnesses “RecentIllness”	Yes No N = 1,685 Missing = 0.41%	421 (24.99%) 1,264 (75.01%)	NA	Logistic	Excluded
Voted in Dec 19 election “Voted2019”	Yes No Not eligible N = 1,676 Missing = 0.95%	729 (43.32%) 879 (52.45%) 71 (4.24%)	Yes vs. others	Logistic	Excluded (fully observed for video)
Party voted “Conservative” “Labour”	Conservative Labour Other N = 682 Missing = 2.76%	300 (43.99%) 297 (43.55%) 85 (12.46%)	Conservative vs. the rest Labour vs. the rest	Logistic	Excluded (fully observed for video)

Self-completion module

Variable and codename	Categories	Observations	Binary transformation	Model	Item non-response analysis
Feeling nervous, anxious or on edge “Anxiety”	Not at all Several days More than half the days Nearly every day N = 1,550 Missing = 8.39%	603 (38.90%) 551 (35.55%) 206 (13.29%) 190 (12.26%)	Not at all vs. the rest	Logistic	Yes
Not being able to stop or control worrying “Worrying”	Not at all Several days More than half the days Nearly every day N = 1,576 Missing = 6.86%	715 (45.37%) 498 (31.60%) 193 (12.25%) 170 (10.79%)	Not at all vs. the rest	Logistic	Yes
Little interest or pleasure in doing things “NoPleasure”	Not at all Several days More than half the days Nearly every day N = 1,585 Missing = 6.32%	845 (53.31%) 423 (26.69%) 185 (11.67%) 132 (8.33%)	Not at all vs. the rest	Logistic	Yes
Feeling down, depressed or hopeless “Depressed”	Not at all Several days More than half the days Nearly every day N = 1,589 Missing = 6.09%	898 (56.51%) 421 (26.49%) 158 (9.94%) 112 (7.05%)	Not at all vs. rest	Logistic	Yes
How satisfied with life nowadays “LifeSatisfaction”	0- Not at all 1 2 3 4 5 6 7 8 9 10- Completely N = 1,692	12 (0.71%) 11 (0.65%) 34 (2.01%) 57 (3.37%) 89 (5.26%) 182 (10.76%) 202 (11.94%) 399 (23.58%) 413 (24.41%) 185 (10.93%) 108 (6.38%)	NA	Linear	Excluded
Financial difficulties in childhood “ChildDifficultiesF”	Yes No N = 1,347 Missing = 20.39%	534 (39.64%) 813 (60.36%)	NA	Logistic	Yes
Number of schools went to by age 16 “NofSchools”	1 2 3 4 5 6	146 (9.01%) 737 (45.49%) 509 (31.42%) 161 (9.94%) 36 (2.22%) 16 (0.99%)	NA	Poisson	Yes

	7 8 9 10 N = 1,620 Missing = 4.26%	9 (0.56%) 1 (0.06%) 1 (0.06%) 4 (0.25%)			
Smoke and how regularly "Smoke"	Never Used to smoke I smoke occasionally I smoke daily N = 1,692	763 (45.09%) 507 (29.96%) 168 (9.93%) 254 (15.01%)	No smoker (1 and 2) vs. Smoker (3 and 4)	Logistic	Excluded
How often has had alcoholic drinks "DrinksRegularly"	Never Monthly or less 2-4 times a month 2-3 times a week 4 or more times a week N = 1,691 Missing = 0.06%	291 (17.21%) 563 (33.29%) 487 (28.80%) 293 (17.33%) 57 (3.37%)	Monthly or less & never vs the rest	Logistic	Excluded
Units of alcohol in a typical day of drinking "AlcoholUnits"	1-2 drinks 3-4 drinks 5-6 drinks 7-9 drinks 10 or more N = 1,399 Missing = 0.06%	495 (35.38%) 475 (33.95%) 287 (20.51%) 90 (6.43%) 52 (3.72%)	NA	Poisson	Excluded
How often drank more than six units in one occasion during the past year "Alcohol6Units"	Never Less than monthly Monthly Weekly Daily N = 1,398 Missing = 0.01%	211 (12.47%) 754 (44.03%) 319 (18.85%) 115 (6.80%) 8 (0.47%)	Monthly or less & never vs the rest	Logistic	Excluded
Financial literacy test #1 – Inflation "FinancialL1"	More Less The same N = 1,683 Missing = 0.53%	320 (19.0%) 1,155 (68.6%) 208 (12.4%)	Correct answer vs. incorrect	Logistic	Excluded
Financial literacy test #2 – Interest "FinancialL2"	Cont. 1 to 130500 N = 1,672 Missing = 1.18%	-	Correct answer vs. incorrect	Logistic	Excluded
Financial literacy test #3 – Interest "FinancialL3"	More than 110 Exactly 110 Less than 110 Impossible to tell N = 1,680 Missing = 0.71%	794 (47.26%) 657 (39.11%) 132 (7.86%) 97 (5.77%)	Correct answer vs. incorrect	Logistic	Excluded