

# Veterans' Survey methodology

An overview of the background, methodology, data collection, processing and weighting of the UK Veterans' Survey 2022.

Contact:  
Emily Knipe  
veterans@ons.gov.uk  
+44 1329 444591

Release date:  
15 December 2023

Next release:  
To be announced

## Table of contents

1. [Introduction](#)
2. [Sample](#)
3. [Survey development](#)
4. [Communication toolkit](#)
5. [Operational delivery](#)
6. [Ethics](#)
7. [Data cleaning](#)
8. [Data processing](#)
9. [Uncertainty and weighting for the analysis](#)
10. [Weighting processes](#)
11. [Presenting results](#)
12. [Cite this methodology](#)

# 1 . Introduction

## Project overview

The Office for Veterans' Affairs (OVA) commissioned the Office for National Statistics (ONS) to carry out the Veterans' Survey to support the aims of the [Strategy for Our Veterans](#), which is to make the UK the best place to be a veteran by 2028.

Stakeholders on the project included the Welsh Government, Scottish Government and Northern Ireland Veterans' Support Office (NIVSO). Charities and veterans' organisations such as COBSEO, The Confederation of Service Charities, supported the project.

## Background rationale

The UK Veterans' Survey was the first project of its kind in the ONS. Although the ONS has conducted surveys in relation to veterans, this was the first time a respondent-driven survey of this size was run, which targeted participants across the UK. The OVA commissioned the research to understand the unique experiences of the UK veteran population and their families living in the UK. The information collected will help inform future policy. The aim of these policies is to improve the lives of veterans and their families across the UK.

## Versions of the survey

There were four versions of the Veterans' Survey, these were:

- the main survey aimed at UK veterans who left the armed forces, were living in the UK and aged over 18 years
- the Veterans' Family Survey for immediate family members of UK veterans, with respondents also residing in the UK and aged over 18 years; the family members could complete the survey if their veteran family member was deceased
- the paper survey of the Veterans' Survey designed to support those who did not have access to technology, such as homeless veterans or disabled veterans who could seek support in completing the paper survey
- the Veterans in Prison Survey for UK veterans in UK prisons was administered in paper survey form

# 2 . Sample

## Sample frame

A voluntary self-select, respondent-driven sample was used for the Veterans' Survey. Prior to settling on this approach, the sampling frame was planned to be sourced from either: random sampling via the Postcode Address File (PAF), random sampling of UK locations with a high density of veterans or from using information from charities, veterans' organisations, the Office for Veterans' Affairs (OVA) or the census to attempt to target addresses of veterans. These considerations were rejected in favour of a self-select sample as the previous strategies had high risks of targeting households that did not have veterans, incurring additional costs, and low response because of missing veteran households.

The respondent-driven sampling method was decided upon because it would allow all UK veterans and their family members to potentially be part of the sampling frame. This would allow us to weight survey responses if needed using the assumption that all veterans and veteran family members aged 18 years and over were eligible to answer the survey (Eligibility decisions are detailed further in our [Ethics section](#)). Therefore, weights created could be considered non-response or final weights. The self-select method was found to be the most cost-effective and time-effective method as it relied upon advertising the survey through targeted communications with veterans' organisations and charities. This reduced costs within the project and reduced risks associated with inviting people that were not eligible.

The possible disadvantages of self-select sampling were considered and included the potential for respondents to not be representative of the known veteran population. To understand this bias, it was proposed that survey responses would be linked to Census 2021 data (for England and Wales) to enable a comparison of veterans' demographic qualities and assess for the sample's representativeness so that re-sampling or adjustments could be made to account for the bias in the analysis. See our [Uncertainty and weighting for the analysis section](#) for further detail. Self-selection bias was also noted as a concern, which is where participants with extreme views or particularly strong experiences may be more likely to participate, however, as studies are voluntary across the Office for National Statistics (ONS), these risks have previously been identified and accepted.

The survey was voluntary for all veterans and family members including veterans who were in prison. Veterans were given the opportunity to withdraw their data from the survey up until the data were combined for analysis. Clear instruction was provided in digital and hard copy with how to contact the ONS to request withdrawal of data. A Welsh language version of the survey was also created.

## Response target

Initial research, to learn from other surveys conducted within the veteran community, was conducted by the OVA to help inform their initial planning for the survey. The ONS helped develop a target response rate based on the estimate of veterans currently residing in the UK, which was estimated to be 2.4 million veterans. A target response rate was created based on 0.5% of this population, or 13,000 veteran participants.

## Collection dates

The survey went live online on Thursday 10 November 2022 and ran for 12 weeks until Friday 3 February 2023. During January 2023, veterans in prison received a paper copy of the survey. The deadline for all paper versions of the survey was Friday 24 February 2023. A week extension was given to assist prison wardens to return batches of paper surveys, which had launched once the main survey was live, and because there was a surge in paper requests in the final week of the online survey.

## 3 . Survey development

# Questionnaire

## Questionnaire specification

The Office for Veterans' Affairs (OVA) explored topics to be included in the survey with government departments, including the Ministry of Defence (MOD), local authorities and charity representatives.

These topics were developed in partnership with the Office for National Statistics (ONS) and with contributions from representatives from the devolved governments. The final set of topics covered a range of issues unique to veterans, ranging from their time in the armed forces, their transition to civilian life and their access to veterans' services and support. Topics also explored their lifestyle, wellbeing, employment and health.

The family members of veterans' have unique experiences in comparison with other families outside of the armed forces. These unique experiences were also explored within the project with a Veterans' Family Survey. This smaller sub-project tied into the main survey with a separate set of topics to be explored from the perspectives of immediate family members of UK veterans.

To allow veterans who could not access the online survey to take part, a [paper survey](#) (PDF, 812.8KB) was created and made available upon request. The [paper survey](#) (PDF, 812.8KB) included the same topics as the main survey tailored to a paper questionnaire design. The paper survey was available in large print and in braille if requested.

The Ministry of Justice (MOJ) requested the inclusion of veterans in prison, using a variation of the paper survey. It was agreed to include veterans who were in custody across the UK, in male and female, private and public prisons. A process was agreed to enable paper surveys to be distributed to prisons and completed by veterans in prison and then returned securely to the ONS. The Veterans in Prison Survey covered most of the topics in the main survey but adjusted for their circumstances, for example, employment options, and access to additional veterans' services.

The survey questions were written and developed between the OVA and ONS using existing questions of interest and significance from Census 2021. Harmonised questions were incorporated, which included standard demographic questions covering date of birth, sex, nationality, sexual orientation, ethnicity and religion. As demographic questions vary according to country within the UK, the main online survey was adapted to adhere to these geographical differences dependent upon the respondent's location. The differences included the ordering of nationalities, religions, and ethnicities. Each of these were routed based on where the respondent lived.

To simplify and condense the paper versions of the survey, which included the veterans in prison survey, the English format for these questions and responses was used. This reduced the length of the paper survey considerably, which reduced the respondent burden. This was considered essential as the paper survey was designed for those who had additional or specific needs, such as disabled veterans, including those who were considered partially sighted or blind.

Harmonised questions were taken from existing surveys from the ONS portfolio and either adapted to be veteran-specific or used as is. Questions on discrimination and bullying were developed to focus on a veterans' circumstances, and new questions were developed to explore new topics of interest including those focused on experiences of active service in the armed forces.

During the iterative question and responses development phase, there was input from stakeholders and teams across the ONS including methodology. Once the questions were established, main topics, phrases and concepts were cognitively tested with a group of 12 volunteer veterans with a range of experiences and backgrounds in the armed forces. These veterans were sourced from the devolved governments, the OVA, MOD, and the ONS. Four females and eight males, aged between 28 years and 64 years took part in the cognitive interviews; all participants identified with the White British ethnic group.

Cognitive testing found that main elements of the questionnaire needed to be adapted because the present-day context of bullying and victimisation was a complex topic to assess when framed within the standard daily experiences of armed forces life. Many veterans stated that, what would now be considered bullying was once considered "character building" behaviour.

The cognitive testing also highlighted the complexities around discussing "active combat", and issues such as security and definition. Questions which were found to segregate armed forces personnel into separate categories were met with resistance. Throughout the discussions with the veteran volunteers, there was a unified belief that all veterans, regardless of sex, ethnicity or sexual orientation were seen as equal and any adaption to that idea was offensive. The specific question which was explored was:

"How can services for veterans today be made more accessible and inclusive for:

- lesbian, gay, bisexual and transgender veterans
- female veterans
- Black, Asian and minority ethnic veterans
- and non-UK veterans?"

After consultation, this question was updated to:

"How can we make services for all veterans more accessible?"

Improvements were made to the wording, structure and topic of questions which were then used as the foundation for the development of the different versions of the Veterans' Survey.

Once the questions were fully prepared and reviewed by the ONS and all stakeholders, a final draft of the question specifications was created. This was then translated into Welsh. The main Veterans' Survey was available in English and Welsh. The Veterans' Family Survey, the Veterans' Survey on paper and the Veterans in Prison Survey were only available in English.

## Eligibility

The eligibility for the survey was set out in the initial planning stages as the project considered ethical implications. It was decided that the Veterans' Survey would be open to all UK veterans currently living in the UK aged 18 years and over. [A veteran was defined as anyone who had served at least one day in the regular or reserve armed forces.](#) Although consideration was given to include all veterans over the age of 16 years, this was rejected because of additional ethical concerns for those under the age of 18 years and because the proportion of veterans who were under 18 years who were likely to answer the survey was considered relatively small and a unique population that would warrant its own study.

Family members, over the age of 18 years, were invited to answer the survey. The relatives' veteran could be deceased or living, and the survey was open to any immediate family member including stepparent and adult stepchildren. Veterans who were in UK prisons were eligible for the study if they were a UK veteran and aged 18 years or over.

UK veterans residing outside of the UK were not eligible for the survey. This was considered the most appropriate course of action as UK legislation would not directly impact this population.

There were concerns about how to ensure those who accessed the survey were genuine veterans and to how to keep access to the survey secure. Password entry was rejected because of logistical reasons and the sampling method. It was suggested that veterans completing the survey could enter their service number to validate they were genuine veterans of the UK armed forces. After consideration this was rejected because of the lack of uniformity across service numbers and that they can alter over the career of an individual meaning that they would not be able to be verified. It was agreed that the risk of false responses was low.

## Mode

The Veterans' Survey would be online-first using [SmartSurvey software](#).

SmartSurvey could be accessed via an online link, using a short URL, and QR code, which allowed ease of access to the survey. These were then added to any digital communications and the communication toolkit. This was the first ONS survey of its kind to use this technology.

The survey was built on the SmartSurvey platform by survey teams within the ONS. It was tested by these teams and representatives of the devolved governments (in English and Welsh). The survey was also reviewed and tested by the Cyber Security team to ensure that SmartSurvey was a safe platform to store data from veterans and to ensure there was minimum risk of sabotage or theft of data.

The paper survey was designed to allow those with no access to technology, lack of confidence with technology and those with disabilities the same opportunity to complete the survey. This was essential as the UK veteran population has an older age structure. The paper survey was only available in English based on demand and costs and only covered the main survey (not the Family Survey). The paper survey was available on request in large print and braille.

The paper survey design was developed using the Census 2021 format. Census demographic questions had received rigorous testing; colours, layout and style were copied almost identically. The paper survey was reviewed and updated by survey teams within the ONS who streamlined and simplified the survey to make the questions, responses and routing most appropriate to a paper mode of collection.

The Veterans in Prison Survey was developed from the paper survey.

All respondents completed the initial eligibility questions, and if the person said they were a family member, they were routed into the Family Survey.

## Veterans in prison

A research request was submitted to the Research Advisory Committee to survey UK veterans within UK prisons. The request process was completed while the survey operation was live. The request was submitted, clarified and approved to proceed in a few weeks. The approval was relatively straightforward as the approach using paper surveys meant no external researchers would be attending prisons to collect the data.

A process flow was developed, which shows how the Prison Survey would be distributed and collected. Stakeholders in all four countries agreed roles and responsibilities within the prison system. New materials, including the survey, invitation, leaflet, privacy notice and envelope were sent as PDFs to MoJ representatives in each country who printed and managed the process locally.

## 4 . Communication toolkit

Participants were recruited to take part in the survey through a communication strategy, which advertised the survey throughout the UK on social media, through veterans' charities and organisations (primarily through the Confederation of Service Charities (COBSEO), via stakeholders and Office for National Statistics (ONS) communications and news outlets.

A brand was created for the survey, utilising the existing style of the Office for Veterans' Affairs (OVA). Digital designs were created, which showed an array of authentic veterans and armed forces personnel as well as family members of veterans. The digital designs also contained the branding of the ONS, OVA, Welsh Government, Scottish Government and Northern Ireland Veterans Support Office (NIVSO).

Slogans were developed and these were used and evolved throughout the communication campaign to raise awareness of the survey. Digital designs were deliberately created to be shared across social media as well as to be incorporated into websites, digital newsletters and emails. Digital designs were also created so prints were suitable quality up to size A3. This allowed organisations the option to print and display posters. Boilerplates were also created to allow stakeholders the ability to use ready crafted content on the survey to allow both ease of drafting communication and uniformity of message being shared on the survey.

While the survey was live, management information was monitored to help feed intelligence into the targeting of the communications. For example, targeting certain ethnic groups.

The materials which stakeholders needed were contained within a communication toolkit, which was hosted on a private webpage. This allowed stakeholders to access materials as and when they needed and when they updated, without the ONS needing to contact separate stakeholders, reducing time and resources.

## 5 . Operational delivery

### Data collection

Data were captured within SmartSurvey and downloaded into secure networks within the Office for National Statistics (ONS) with limited access to ensure security and confidentiality of data. The data were then incorporated into the Data Access Platform (DAP), an integrated digital platform, which provides the tools and technology needed to store, analyse and process data in the ONS. Both paper surveys were designed to be able to be scanned into digital readers in the ONS to allow for the paper surveys to be digitalised.

### Information shared with veterans

To ensure that veterans and family members were provided with information regarding the aims of the survey, confidentiality, and General Data Protection Regulation (GDPR) statements, a webpage was created for the online survey and paper materials were made for the paper surveys. It also gave a brief introduction to the survey and signposted to additional support services, specifically the Veterans' Gateway.

### Accessing the survey

The survey was designed to allow participants the opportunity to save the survey and continue later. Participants could add their email address while completing the survey and exit the survey part way through. Email instructions were then sent to the participant to allow them to continue where they left off.

This was an essential element incorporated into the survey completion process, which allowed the survey to be completed over multiple sittings. This was supportive to a population which is predominantly older and has a proportion who have health, disability and lifestyle vulnerabilities.

## Help with completing the survey

Veterans were offered support to help complete the survey through multiple avenues. A dedicated mailbox allowed veterans to communicate directly with the ONS and the Survey Enquiry Line offered over-the-telephone assistance in completing online or in requesting a paper copy of the survey. Several stakeholders and organisations provided direct support, either by taking printed copies of the survey to veterans in their own homes, allowing veterans to access technology in their offices or with support workers, or by assisting them with digital and paper completion as needed. Libraries across the UK were also informed about the survey and offered both access to technology and one-to-one help with completing the survey online.

If veterans wanted to take part in the survey but were unable to complete it themselves, it was possible for someone to answer on their behalf as a proxy. In these instances, personal information was not requested as the participant may not want to share those details with an assistant. Routing was included in the survey to stop those answering as proxies from being asked personal questions (questions from sexual orientation to bullying and harassment were skipped in these instances).

During the live data collection period, updates were made to the survey responses for some of the questions. These were deemed necessary as veterans and representatives from various veteran groups contacted the ONS with concerns regarding listings of ranks within the armed forces and with requests to update their status to veterans as listed within our qualifying services. Terminology was also requested to be updated regarding leaving service. The ONS along with the Office for Veterans' Affairs (OVA) agreed the updates and altered the lists of ranks to include, for example, "Master Aircrew" in the RAF ranks and "Other" within all the ranks to encompass any missing ranks.

After the high level of engagement from veterans directly with the ONS, a separate report and presentation was given to the OVA, which shared the views and concerns veterans had raised in relation to both the survey and topics of general concern.

Veterans who were not currently within the UK would not be eligible for the survey. Although this appeared to be a logical decision in the planning stages, during the live operation of the survey, it was found that UK veterans living outside the UK were keen to take part in the survey. Multiple individuals, groups and representatives of groups from around the globe contacted the ONS requesting to be involved. It was decided not to change the eligibility of the survey based on the scope of policy making concerning veterans and families living within the UK and the disruption changing the live survey would cause to the data collection. This rationale was shared with veterans living outside the UK who had corresponded with the ONS directly.

## MiHub

SmartSurvey provided management information (MI) on the responses and this MI was used to show progress during the data collection operation. The MI over-estimated "complete responses" as it included partial responses, so the two were monitored separately.

The ONS team produced the age profile (which had to be derived from date of birth) during the operation. This showed that fewer older veterans (75 years and over) were participating. The charities and veterans' organisations were asked to focus on those aged 75 years and over and given additional material to encourage older veterans to take part.

## Procurement

The Office for National Statistics (ONS) is the data controller for this project, which means they were responsible for determining the purposes and means of the processing of personal data, in accordance with the General Data Protection Regulation (GDPR).



## 6 . Ethics

An Ethics application was made to the ONS Ethics Committee to consider any ethical implications of the survey. As well as the standard concerns to be investigated, there were unique considerations of the veteran population. Veterans are older and contain vulnerable individuals who could have experienced trauma during their time in service, or difficulties transferring from the armed forces into civilian life. The veteran population is estimated to have individuals with a lower-than-average reading age, and a proportion of the population have complex mental health requirements. This group also has a proportion of individuals vulnerable to addiction and homelessness.

The Ethics application also considered the suitability of the survey for respondents under the age of 18 years. Because of the considerations made during the Ethics application, the survey was tailored to meet the needs of a vulnerable population or reduce risk of harm.

It was recommended that only adults aged 18 years and over complete the survey, or the ONS would need to radically overhaul the questions on active combat and wellbeing. This recommendation was followed as it allowed the topics covered in the survey to remain unchanged; the number of potential veterans under the age of 18 years who would take part in the survey were considered small in comparison. The recommendations also allowed the planning for the survey to be more inclusive and accessible to those with additional needs or without access to technology and to be more respectful by ensuring the most positive terminology was used when referring to veterans in different circumstances, such as in prisons.

### Data protection

The data were stored and managed securely by the ONS. A Data Privacy Impact Assessment (DPIA) was approved by the ONS.

### DPIA

The DPIA set out the aims, benefits and risks of the survey regarding collecting, processing and storing personal data. A huge amount of personal and sensitive data were collected including individuals' date of birth, sexual orientation, ethnicity and other demographic data. Sensitive data on personal experiences during individuals' time in the armed forces and mental and physical health status was also collected.

The DPIA was put in place to ensure the considerations for collecting, processing and storing the data securely had been made and risks had been minimised. The DPIA detailed the steps which would be taken to prevent individuals from being identified and to keep the data safe on SmartSurvey and during processing. The DPIA also detailed the steps taken with stakeholders and other parties within the ONS, to keep the data securely.

## 7 . Data cleaning

Once the survey closed, the data were cleaned prior to being processed. This involved various checks to ensure answers given were logical and valid. This was specific to the paper and prison surveys as routing and valid value rules were built into the online questionnaire.

### Dataset creation

SmartSurvey stored partial and complete responses separately for both the veteran and veterans' family respondents. Each separate questionnaire dataset needed to be extracted from SmartSurvey. To prepare the final dataset four separate datasets were extracted. These were a partial and completed Veterans' Survey and a partial and completed Family Survey.

The dataset for paper and veterans in prisons responses was created separately as these paper surveys needed to be manually entered. UK prisons printed off the surveys themselves, and so the returned paper copies varied significantly meaning that a tool could not be created to scan all paper responses into a dataset.

For the purposes of time, internal ONS volunteers were recruited to support the entry of paper survey responses by inputting them into an Excel spreadsheet, which was then exported into a dataset in Statistical Package for the Social Sciences (SPSS) software to be processed. Once all the datasets were extracted, variable names and labels were added to all variables, which allowed the datasets to be added together to create one dataset containing every survey respondent.

## Routing checks

Routing statements and instructions were included throughout the paper questionnaire, however, there were instances in which questions were answered despite the survey not routing to them. To clean this, all variables which contained or were impacted by routing statements had routing checks written to determine whether any questions that were not routed had been answered.

In instances where this happened, these were examined to determine whether the answer was logical or not, to decide whether the value should be overwritten with a minus 9 or if other variables should be changed to enable it to be routed. For instance, some veterans answered "no" to whether they were registered with a dentist prior to continuing to answer questions about whether they were registered to an NHS or a private dentist. As logically this suggests the respondent is registered, the answer to the original question about dentist registration was changed to "yes". If they were not registered, the answer to the type of dentist would be overwritten with a minus 9 as they should not be routed.

## Data validation

Further cleaning was required in instances where respondents selected multiple answers in a question which required only one answer. In the paper survey only multi-select questions had guidance for how to answer the questions, the other questions did not specify that only one answer was expected.

An example of this is that there were some instances in which respondents selected multiple answers for the rank question "What was your rank upon leaving the [Service Name]?". From an analysis perspective, the rank variables were used to inform the derivation of an overarching rank-derived variable which showed officer and non-officer ranks. As a result of this it was decided that the highest rank would be used to ensure that anyone that held an officer rank was captured in the derived variable (DV).

There were other instances in which multiple answers were selected; these were explored on a case-by-case basis and some assumptions were made to ensure the answers were logical.

Questions which included an "other" option were also investigated as part of data validation checks to ensure that only true "other" values were retained and others, if applicable, were put into the right category. This was done for each rank question where there were various "other" answers that applied to one of the listed categories. For instance, in OtherRAFRank answers included "Senior Aircraftsman" and "Senior Aircraft Woman", which were covered in RAFRank = "Junior Technician/Leading Aircraftman/Senior Aircraftman". In this case the OtherRAFRank would be changed to minus 9 (not routed to) and RAFRank would be overwritten with "Junior Technician/Leading Aircraftman/Senior Aircraftman".

## Veterans in prison

As part of the cleaning process, prison responses required special attention to input postcodes and addresses based on the prison which they were returned from as well as there being some abnormalities in the responses.

To support census linkage, prison addresses were added to the prison survey responses. There was a checklist sent out when surveys were sent to prisons, which was then returned with each prison's batch of surveys that included information such as which prison the return was from as well as the number of veterans that completed the survey. This often included the prison address, which was then applied to every veteran in that batch as they could be identified by the survey ID.

In some cases, the prison checklist was missing. To identify the prisons that these were sent from, the surveys were looked at to see if there was a mention of the prison or if someone had provided a postcode when asked. In this instance we would then be able to find the prison address and apply it to every survey that was returned from that prison. Of returned prison surveys, 23% were unable to be given a postcode.

Other cleaning was also applied to the veterans in prison responses. Pre-determined rules for answers to employment and housing questions were applied to all veterans in prison. However, duplicate prison-specific versions of some variables were made to ensure that the original responses could be retained while the expected values for prisoners could be used in analysis. The housing situation of prisoners was changed to "other" as opposed to responses that may suggest they were not currently residing in prison.

In addition, employment status was changed to "None of the above" so that veterans in prison are shown as economically inactive for analysis, as expected. Possible reasons for this being answered otherwise include respondents selecting answers that seemed appropriate such as "temporarily away from work" or answering based on any jobs they may have within the prison. In the original dataset, 32% of respondents answered with employment status answers that were not expected, with 13% answering "working as an employee" and 7% answering "doing any other kind of paid work". Once changes had been made using the predetermined rules, all responses to questions these respondents should not have been routed to were changed to minus 9 (not applicable).

There were also instances in which those responding via a prison answered "No" to having been convicted of a criminal offence, having served a prison sentence and to whether they were currently serving a prison sentence. As they answered via the Veterans in Prison Survey, all these answers were changed to "Yes" for consistency, although there will be some circumstances where this is not the case.

## String variables

Special attention was required when cleaning free-text responses as these enabled participants to write unlimited without any validation rules in SmartSurvey. These did not have a word or character limit meaning that response length varied.

As respondents could write anything in these boxes, every answer had to be checked to identify and redact disclosive information such as names, addresses and dates. In some instances, very specific scenarios were explained in depth, which needed to be heavily censored to prevent a response from being disclosive.

When checking the data for disclosive information, the ONS also checked for any possible safeguarding concerns if a respondent appeared to be a danger to themselves or anyone else. The ONS safeguarding policy was followed and in some instances, responses were flagged and raised with a safeguarding officer to investigate, resulting in services such as the [Veterans' Gateway](#) being signposted for support.

## 8 . Data processing

Processing of the data included preparing the dataset, replacing missing values, routing checks, removing duplicate cases, creating derived variables, and checking free-text responses for safeguarding or disclosive responses.

### Derived variables

Several new variables were created as part of processing, which aid analysis. Derived variables (DV) were created for ethnicity to group responses together, an age variable was created, and a rank variable was created to identify officer and non-officer ranks.

## DVEthnic

The ethnicity DV (DVEthnic) was created to group ethnicity variables making them comparable with the census groups. In the online version of the survey the ethnicity variable was different based on the answer to PartUK (In which part of the UK do you live?). Those that answered part UK = 1 (England) were asked EngEthnic, the same applied for the remaining answers for PartUK, which related to either Scotland, Wales or Northern Ireland. Each individual granular ethnicity was mapped to one of the five high-level census groups, these are:

- Asian, Asian British, Asian Welsh
- Black, Black British, Black Welsh, Caribbean or African
- Mixed or Multiple
- White
- Other ethnic group

Ethnicity was collected in just one variable in the paper survey, the responses were all mapped to the high-level groups. This was because of there being more complex routing online and to reduce the number of questions in the paper version of the survey. As a result of this, there were some differences in the options in the survey response groups, with the online version providing more granular options for each of the devolved nations. Ethnicity was not captured within the Family Survey.

## DVAge

Age was not collected in the survey, so the date of birth was used to derive the age of respondents. To do this, the date of birth was taken away from the date that a respondent ended the survey (Ended).

For the paper survey, there was not an automatic value in the Ended variable, in this case the date that the paper survey closed was applied. The logic was the same for the calculation of age for those that answered the Family Survey. In some instances, respondents did not have a complete date of birth; the year of birth (DOBYear) was then used as an alternative to ensure that, where possible, respondents' ages were retained.

## DVRank

A veteran-specific DV was created for rank. This was not applicable to the Family Survey. This DV was designed to differentiate between officer and non-officer ranks. The Office for National Statistics (ONS) worked with the Ministry of Defence (MOD) and Office for Veterans' Affairs (OVA) to group the list of ranks for each service type into these categories.

"Other" rank responses were given an "other" value here as they could not be allocated to either category.

## Flag variables

Flag variables were created to help identify cases with certain characteristics, these include an eligibility flag.

## Eligibility flag

An eligibility flag was used to determine who had answered the survey and was eligible. For the main Veterans' Survey (online and paper) and the Veterans in Prison Survey, the predetermined criteria of respondents being UK veterans currently living in the UK aged 18 years and over was used to inform this.

Firstly, everyone was given an eligibility flag of 0, showing everyone as eligible before overwriting this with a 1 if they did not fulfil the criteria. This was done iteratively to ensure that only eligible veterans were retained.

Responses to the question "Which part of the UK do you live in?" were used to determine whether veterans currently lived in the UK. This excluded those that answered "I do not live in the UK" and anyone that had a missing response for this question. Missing responses were considered ineligible as there was no way to verify the respondent's eligibility.

DVAge was used to identify whether the respondent was aged over 18 years. The final part of the eligibility flag was to ensure that no one that is currently serving was eligible, and that respondents had all previously served in the UK armed forces.

To ensure that everyone that had previously served was retained, responses to the question which asked about previous service in the UK armed forces was used. Any respondent not stating they had previously served in the regular or reserve armed forces or that they had previously served in national service was made ineligible. Missing responses were made ineligible as well as those that answered, "none of the above". To remove any respondents that are currently serving, only respondents who said they were not a current serving member of the regular or reserve UK armed forces were considered eligible. This means that missing responses as well as those that said they were currently serving were made ineligible.

For the Veteran's Family Survey, all respondents were considered eligible. One of the criteria for answering the family member question was for respondents to be aged over 18 years.

Because of a high level of missingness in the family date of birth variable (FamDOB), a high proportion of family respondents would be considered ineligible. Given there were explicit instructions at the start of the survey outlining a minimum age for respondents, it was decided that all of these would be retained initially, and further rules would be applied in analysis to ensure that only veterans' family members that we cannot be sure are aged under 18 years will be excluded. This will be done using a set of rules outlined in future analysis of the Family Survey based on responses to other questions respondents have answered. For example, if a respondent has stated they are a grandparent of a veteran, we can assume they are likely to be over 18 years and so would be eligible.

## Completion status flag

For completion status, the flag was created to show those that had been present in the initial "complete" extract of the Veterans' and the Family Survey. To maximise the amount of data that could be used for analysis there was an investigation into how much of the survey had been answered in the "partial" dataset and whether any cases could be retained, in which case the "compstat" variable would be changed from "partial" to "complete".

It was determined that anyone who answered at least one of the final 10 questions as well as either sexual orientation or gender identity could be included as "complete" as they answered most of the survey including some demographic questions.

## Completion Mode

A completion mode flag was created to determine how people responded to the survey. This shows whether the survey was accessed online, via prison or completed on paper.

To derive this, all cases uploaded from SmartSurvey were given a value of completionmode = 1 (Online). To differentiate between paper and prison responses, the SurveyID was used and veterans in prison SurveyID's began with "VPR" while other paper survey IDs began with "VET" or "VLP" (Large print). The prefix was used when inputting paper data entries to give completionmode = 2 for prison completions and all other paper surveys having completionmode = 3.

## Missing responses

If respondents had not answered any individual questions, because of either skipping a question on purpose or being routed away from a question, these then showed as "system missing" in SPSS. These were given a default code for "not answered" (minus 8).

To determine between these two types of "not answered", routing checks were applied to code any responses that should not have been answered by a respondent (minus 9). To do this, routing statements were used to identify any variables which were not due to be answered. This involved writing syntax containing the inverse of the routing to select any answers which should have been skipped. This was done for every variable which was impacted by routing and checks were carried out to determine whether a valid value or a minus 9 (not applicable) should be present.

## Duplicates

As the survey was self-select, there was a risk of people being able to fill in the survey multiple times by either completing the survey on different devices or just accessing the survey again. Only one entry per respondent could be used and duplicates needed to be identified and removed.

We used SPSS' "identify duplicate cases" function to identify which cases had duplicate first names, surnames and date of births as well as duplicate IP addresses. From these, the latest version of the data was retained (these were selected using the "Ended" variable, which indicated when the survey was last accessed) and the others were given a duplicate flag. The flag variable was then merged on to the full dataset and all cases which had been flagged as duplicates were removed. In total 2,533 cases were removed.

# 9 . Uncertainty and weighting for the analysis

## England and Wales veteran responses

For England and Wales survey responses, we have linked the survey data to Census 2021 data and have also compared survey respondent profiles with veteran profiles from Census 2021 to understand how representative we think the survey was of the known veteran population. This aids interpretation of the survey data and has informed weighting decisions where bias in the survey profile was apparent, including the assumptions we have made for the other UK countries based on the findings of how representative the England and Wales responses were.

We have also published a [further report](#) explaining the process behind linking the survey data to Census 2021 data. This covers the quality of the linkage and any bias discovered in which survey respondents were linked and which survey respondents were not linked. This bias was considered in our interpretation of how representative the survey was and should be considered if analysis of the linked data only is conducted.

Census 2021 gives us a robust understanding of the profile of veterans living in England and Wales as of March 2021. We are satisfied that the time lag between Census 2021 and the Veterans' Survey 2022 (20 months) will not have impacted the demographic profile of veterans to a degree that means we cannot use Census 2021 as a population base for veterans in England and Wales. We therefore used Census 2021:

- to assess any bias in the demographic or personal characteristics of veterans that responded to the survey
- to adjust via raking weights (where appropriate) the England and Wales survey data to make them more representative of the veteran population from Census 2021

## Weighting England and Wales Veterans' Survey data

The age profile of veterans responding to the survey differed markedly to the age profile of veterans identified in Census 2021, with survey respondents being younger than veterans from Census 2021. This may reflect the fact the survey was predominantly online and/or that marketing and promotion of the survey was more likely to reach younger veterans.

We used raking techniques to generate weights for England and Wales survey responses, based on the proportions of veterans we would expect to be within given age bands when we considered the age range of veterans from Census 2021. Age bands assessed were ages 18 to 24 years, and then increasing five-year aged bands up to age 90 years (for example, age 25 to 29, age 30 to 34, and so on). The final age band was age 90 years and over. As all veterans and veteran family members aged 18 years and over from England and Wales were eligible to answer the survey, weights created could be considered non-response or final weights. The weights were an adjustment for age distribution only. Weights were scaled to sum to the number of respondents from England and Wales, such that the weights had an average of 1.

## Northern Ireland and Scotland veteran responses

There are currently no veteran data available for Northern Ireland and Scotland close enough by time lag that they can be reliably used to assess the representativeness of responses to the survey from these countries.

For this reason, analysis at country level for Northern Ireland and Scotland will be unweighted. For UK level analysis all respondents that stated they lived in Scotland or Northern Ireland have been allocated a weight of 1. This principle was maintained even when a respondent gave a postcode that suggested they had an alternative address in England or Wales. However, some assumptions will be made about bias in respondent profiles from Northern Ireland or Scotland, based on biases we identify in the survey respondents' profiles from England and Wales as compared with data from Census 2021, which gives us a robust understanding of the veteran population in England and Wales.

## Family Survey responses

2,390 people responded to the Family Survey and response rates varied by question asked. We are aware this represents a small proportion of family members that would have been eligible to respond to the survey. These data will be presented as unweighted data only.

## Measuring uncertainty: UK, Northern Ireland and Scotland

Standard errors "assuming a simple random sample" account for the survey sample size and the variability in the sample responses, but not additional variability associated with the design. It can sometimes be appropriate to publish these simple random sample standard errors, but when calculating statistics based on complex surveys, it is normally more appropriate to use standard errors that consider the complex design, which can be approximated by design factors. This enables understanding of the "complex" or "true" standard error, as described in [ONS methodology working paper series number 9 – Guide to calculating standard errors for ONS Social Surveys](#).

Weights have not been applied to veteran responses from Northern Ireland or Scotland; we have assumed no bias by age for these data and respondents only represent themselves. No age profile data are available for veterans in Northern Ireland. However, there is some evidence that [UK armed forces veterans in Scotland had a similar age profile to those in England and Wales historically](#).

For UK-level analysis, design effects taken from England and Wales weighted data have been applied to standard errors calculated from UK-level estimates and sample sizes, assuming a simple random sample.

## Measuring uncertainty, England and Wales

Since all England and Wales data for veterans have been weighted, uncertainty will be calculated using R statistical software to take into account additional variability introduced through weighting, according to the principles outlined in section 3.4 of [ONS methodology working paper series number 9 - Guide to calculating standard errors for ONS Social Surveys](#).

## 10 . Weighting processes

We found differences between Census 2021 veterans that linked to a survey response from England and Wales compared with veterans who did not. In the main, these findings were replicated when comparing veterans responding to the survey from England and Wales and veterans within Census 2021 by personal characteristics, suggesting these were a result of bias in sample profile. This is outlined further in [The Veterans' Survey 2022. demographic overview and coverage analysis. UK](#).

In selecting which if any variables to include in the generation of raking weights for veterans from England and Wales, it was important to balance between reducing bias in the survey results and introducing additional uncertainty in our estimates based on variability introduced by the implied survey design because of weighting.

### Age

The largest bias identified was age. We found survey respondents in England and Wales to be substantially younger than would be expected, if we used Census 2021 as a population base for veterans in England and Wales.

We used the [anesrake](#) package in R to assess how far the age bands of survey respondents differed to census respondents and found large differences, which justified raking the survey data using age band proportions for veterans from Census 2021 as a target. Age is also strongly related to several outcomes that might be of policy interest in veterans research (for example, health and employment status). Given the large differences in age between Census 2021 and the survey, and the known relationship between age and outcomes of interest, age was selected as a raking variable.

### Sex

Smaller differences between the survey and the census were noted in relation to sex. We found sex to differ by fewer than 5 percentage points between Census 2021 and the England and Wales survey responses from veterans. In addition, the survey attracted few veterans that has served only as a reserve and females were more likely to have served as a reserve only than as a regular member of the armed forces based on data from Census 2021. When we considered veterans that has served as a regular only, the difference in the sex ratio of Census veterans compared to the survey respondents, became even smaller (less than 0.2 %). Because of this we decided that bias identified was not large enough to justify any additional uncertainty that would have been introduced into our findings if we also included sex as a raking variable, regardless of this being small.

### Region

Large differences by region were noted with the North West region, being underrepresented by survey respondents as compared with Census 2021. However, some survey respondents (5.18%) did not provide a postcode or did not provide a postcode that could be successfully attached to a region. This means that creating raking weights for England and Wales survey respondents based on region is not feasible without additional imputation.

### Economic activity

We found differences between Census 2021 and responses from England and Wales to the survey in relation to economic activity, with the survey having proportionally more people employed and consequently fewer people retired. However, once we weighted the survey responses for England and Wales by age band based on Census 2021 proportions, this difference became much smaller, suggesting this finding was driven by the age bias already addressed.

### Other characteristics

[The Veterans' Survey 2022. demographic overview and coverage analysis. UK](#), discusses other characteristics where we expect a small bias in the sample profile that would not be suitable for raking purposes. These should be considered when interpreting findings from the survey.



## Process

As age band was selected as the one variable for which we would weight the England and Wales survey data to be more representative of Census 2021. We used the standardised approach for developing weights using raking (rim weighting or iterative poststratification) as described in [Standardizing and Democratizing Survey Weights: The ANES Weighting System and anesrake \(PDF, 756KB\)](#). This approach reduced the weighting variability as compared with directly applying Census 2021 age band proportions directly to the survey data because we were able to use iteration techniques appropriately to test various caps for the maximum weight.

The presence of very large weights can lead to analyses that are sensitive to the reports of but a few individuals. This latter step resulted in a drop of the maximum weight assigned from almost 13 to 5 and this in turn reduced the variance associated with the weights. This step led to a reduction on the overall design effect (from 2.9 with a cap of 12.5 to 1.9 with a cap of 5). Age groups most affected by this cap were those aged 18 to 24 and those aged 75 years and over. Once we had selected the cap to be applied and the weighting variable of interest, we used the [anesrake](#) package to calculate weights and applied those weights to the England and Wales dataset. Convergence was successfully reached after 4 iterations.

Rather than use this generic design effect to calculate standard errors, we have taken steps to use design effects specific to the analysis being conducted using statistical packages in R, but the impact of not capping the weights would have been similar to that described previously for all design effects considered.

## 11 . Presenting results

### Confidence intervals

Survey estimates are presented with 95% confidence intervals. Any sample survey may produce estimates that differ from the figures that would have been obtained if the whole population had been interviewed. It is, however, possible to calculate a range of values around an estimate, known as the [confidence interval](#) (also referred to as margin of error) of the estimate.

At the 95% confidence level, over many repeats of a survey under the same conditions, one would expect that the confidence interval would contain the true population value 95 times out of 100. Confidence intervals presented are based on complex [standard errors](#) (CSEs) around estimates, which reflect the survey design.

### Robustness

The number of cases upon which analysis is based is important as it influences the precision (standard error) of the estimates. Estimates where the unweighted base is less than 50 cases are not published since the confidence intervals associated with these estimates can be very wide and add little value to interpretation.

## 12 . Cite this methodology

Office for National Statistics (ONS), released 15 December 2023, ONS website, methodology, [Veterans' Survey methodology](#)