Occupation Coding During the Interview in a Web-First Sequential Mixed-Mode Survey

Darina N. Peycheva[1], Joseph W. Sakshaug[2], and Lisa Calderwood[1]

Coding respondent occupation is one of the most challenging aspects of survey data collection. Traditionally performed manually by office coders post-interview, previous research has acknowledged the advantages of coding occupation during the interview, including reducing costs, processing time and coding uncertainties that are more difficult to address post-interview. However, a number of concerns have been raised as well, including the potential for interviewer effects, the challenge of implementing the coding system in a web survey, in which respondents perform the coding procedure themselves, or the feasibility of implementing the same standardized coding system in a mixed-mode self- and interviewer-administered survey. This study sheds light on these issues by presenting an evaluation of a new occupation coding method administered during the interview in a large-scale sequential mixed-mode (web, telephone, face-to-face) cohort study of young adults in the UK. Specifically, we assess the take-up rates of this new coding method across the different modes and report on several other performance measures thought to impact the quality of the collected occupation data. Furthermore, we identify factors that affect the coding of occupation during the interview, including interviewer effects. The results carry several implications for survey practice and directions for future research.

Key words: Occupational classification; self-administration; interviewer-administration; coding error.

1. Introduction

The collection and coding of respondent occupation has been one of the most important, yet challenging, tasks of social surveys for decades. Occupation coding, traditionally performed manually and post-interview, has been acknowledged as time consuming, costly and error-prone. The challenges of manual occupation coding have led to innovations in the use of computer-aided occupation coding, performed by coders or interviewers during or after the interview (Lyberg and Dean 1992). However, with the increasing use of online and mixed-mode surveys, still little is known regarding the feasibility of coding occupation during the interview and factors contributing to the performance of the occupation coding instrument in online and mixed-mode surveys.

[1] Centre for Longitudinal Studies, UCL Institute of Education, 55–59 Gordon Square, London WC1H 0NT, United Kingdom. Emails: d.peycheva@ucl.ac.uk and l.calderwood@ucl.ac.uk
[2] Institute for Employment Research, 104 Regensburger Straße, Nuremberg 90478, Germany. Email: joe.sakshaug@iab.de

This article addresses these issues by presenting results from a large-scale sequential mixed-mode (web, followed by telephone, then face-to-face) cohort study of young adults in the UK, in which respondents were asked to self-code their occupation online in the initially offered web mode, and interviewers were tasked with identifying the relevant occupation code via telephone or face-to-face when an online interview was not possible.

To our knowledge, use of computer-aided occupation coding during the interview in large-scale probability-based surveys is still rare, and evidence on its performance is not readily accessible in the survey literature. Furthermore, the performance of occupation coding during the interview has not been investigated in a web-first sequential mixed-mode survey, nor the extent to which respondents and interviewers influence the use of the coding instrument. However, the importance of these issues has been acknowledged for future improvements of occupation coding instruments (Belloni et al. 2016; Conrad et al. 2016; Schierholz et al. 2018).

In the following sections, we provide a brief overview of the relevant literature (Section 2), outline the research questions (Section 3) and methods used to address them (Section 4), present the results (Section 5) and discuss the conclusions and practical implications of the findings (Section 6).

2.   Literature Review

It is common practice in social surveys to collect information about occupation with a series of open-ended questions asking participants for their job title and to describe the kind of work they do. Such questions enable the collection of sufficient detail about respondents' occupation and assignment of a code at the most detailed level of the occupational classification. These questions may be administered in both interviewer- and self-administered survey settings, and are also asked in mixed-mode surveys. Less frequently, occupation is captured with closed-ended questions offering limited choice of occupational categories, thus resulting in highly aggregated occupation codes. Alternative formats have been offered in web-based surveys, including search tree navigation and semantic text matching techniques, and look-up methods to self-identify one's own occupation. Their application in computer-assisted personal and telephone interviews has also been documented (Tijdens 2014, 2015a). Open-ended occupation questions, however, still dominate the research practice. The verbatim responses collected from these questions are typically converted into occupation codes post-interview by specialist coders using manual or computer-aided (computer-assisted or automated) coding procedures. Today, the use of manual coding has been significantly reduced, but it still complements computer-aided coding methods.

In manual coding, coders assign an occupation code based on the open-ended responses, using a standardized classification scheme without any degree of computer assistance. While classification schemes differ, they all include hundreds of occupation codes nested within hierarchical groups with more specific occupation groups nested within more general groups. For example, the 2010 classification of occupations in the UK, known as the Standard Occupational Classification (SOC2010), and used for occupation coding in the current study, has 9 major (1-digit), 25 sub-major (2-digit), 90 minor (3-digit), and 369 unit (4-digit) (Office for National Statistics 2010a). Manually selecting a code at the most

detailed (4-digit) level of the classification scheme is a time-consuming and expensive process, but also error-prone as even professional coders, following detailed coding guidelines, might disagree on the occupation code for a given case (Lyberg and Dean 1992; Creecy et al. 1992; Campanelli et al. 1997). Manual coding has been found to be especially problematic by National Statistical Offices (NSOs) where coding is extensive (e.g., censuses and large-scale sample surveys), with manual coding error rates of 10% or greater (Lyberg and Dean 1992; Creecy et al. 1992).

Faced with this challenge, the endeavour of computerizing the coding of open-ended responses dates back to the early 1960s, with first applications of computer-aided occupation coding in the late 1970s (Lyberg and Dean 1992; Creecy et al. 1992). Two main forms of computer-aided coding emerged – computer assisted and computer automated. In the former, as in manual coding, the coding is performed by a coder who works interactively with the computer, which guides their decision. In the latter, occupations are coded automatically by software. However, fully automated coding is rare in practice. Rather, it is usually supplemented with manual or computer-assisted systems. Automated coding usually codes part of the occupation entries in which a desired level of certainty associated with the occupation code is set. That is, occupation descriptions with a high degree of certainty (i.e., above a certain threshold) are coded automatically, otherwise human intervention is sought. Even though coder involvement is significantly reduced, human coding is preferred for the so-called "hard-to-code" or "difficult" occupations. This is also known as semi-automated coding, in which automated coding is complemented by human coding for certain situations, compared to fully automated coding, in which 100% of the coding is performed by the software. Computer-aided coding can be administered during or post-data collection, although post-data collection coding is more prevalent.

Computer-aided coding involves computer-stored dictionaries that can be built from coding manuals (e.g., classification schemes) or on empirical patterns of responses provided by respondents in earlier surveys (i.e., previously coded occupation information from previous studies, pilots, etc.), or a combination of manuals and empirical patterns. However, dictionaries constructed from manuals have been considered a disadvantage in that they are strongly dependent on the experience of the coder (e.g., respondents may use words or phrases not included in the manual). In contrast, it has been considered more efficient to base the dictionary on empirical response patterns, in which matching can benefit from the use of words and phrases given by previous respondents (Lyberg and Dean 1992; Creecy et al. 1992).

Various forms of matching have been applied, generally classified as rule-based and data-driven (e.g., statistical or machine learning) (Gweon et al. 2017; Schierholz and Schonlau 2020). For example, if the open-ended answer meets a prespecified logical condition (e.g., presence of a certain word), then a specific code is assigned. Such 'if-then' statements are called 'rules'. Rules are written by experts or can be based on previous data analysis. More recently, statistical learning or machine learning approaches have been employed, whereby a model is trained on manually coded training data and used to predict the most probable code for new data (Gweon et al. 2017; Schierholz et al. 2018; Schierholz and Schonlau 2020).

Evaluations of occupation coding performance between coding methods are scarce, mostly based on comparisons with manual coding used as a 'gold standard', despite the

challenges it poses. Some of these studies have shown that automated coding works reasonably well in reducing the number of cases to be manually coded post-interview, but that it is not yet ready to replace human coders (Ossiander and Milham 2006; Burstyn et al. 2014; Helppie-McFall and Sonnega 2018). Belloni et al. (2016), using the automated coding tool as a benchmark, also stressed the benefits of using automated and human coding jointly.

Campanelli et al. (1997) compared manual coding with computer-assisted and automated coding. They found only a modest gain in performance using computer-assisted compared to manual coding. Automated coding was particularly sensitive to the amount and type of input which was entered, and sometimes scored significantly lower than manual coding with respect to the plausibility of the occupation code. For example, when both the job title and job description were used in the search algorithm, automated coding was comparable to manual coding. However, when the description input was limited and only the job title was used, the quality of coding was substantially lower than manual coding. In terms of time-saving, manual coding and computer-assisted coding did not differ – a result that the authors suggested was due to the fact that the coders were new to the coding software. As expected, automated coding yielded the largest time (and cost) savings, assigning a code in nearly all (99%) of the cases.

The level of detail of the verbal descriptions used in automated coding was also acknowledged by Belloni et al. (2016). The authors used data from the Dutch sample of the Survey of Health, Ageing and Retirement in Europe (SHARE) and compared manually coded verbatim responses on current and last occupation with the codes assigned by the Computer Assisted Structured Coding Tool (CASCOT) occupation coding software. The authors found that automated coding was significantly improved when additional auxiliary information, such as training and qualifications needed for the job and the industry in which the respondent is working, was included in the search algorithm. In contrast, Helppie-McFall and Sonnega (2018) found that the NIOSH Industry and Occupation Computerized Coding System (NIOCCS), employed in the Health and Retirement Survey (HRS) to code occupation history data, and compared to coding results from a highly trained human coder, worked well only with short descriptions, one to three words each, of job title or job description (and industry) as inputs – a finding in line with Conrad et al. (2016).

Conrad et al. (2016) found the length of the occupation description to be a factor strongly related to the reliability of post-interview coded occupations. However, the observed relationship was dependent on the particular occupation terms. For example, for 'easy' occupation terms, longer descriptions were less reliably coded than shorter descriptions, but for 'difficult' occupation terms, longer descriptions were slightly more reliably coded than shorter descriptions. The authors argued that the occupation descriptions do not necessarily need to be long or overly specific, particularly for 'easy' occupations, and that interviewers should rather be trained on the logic and rationale behind the coding structure so that they have a better sense of the kinds of decisions coders need to make. That longer descriptions do not necessarily result in more accurate occupation coding was supported by Massing et al. (2019) who found that reliability decreased as descriptions became longer. Bergmann and Joye (2005) also suggested that the more detailed the information to be coded, the less reliably individual cases are

assigned to categories. Cantor and Esposito (1992) reported that coders who were asked to comment on recordings of interviewers' questioning strategy only rarely indicated that more detailed information would be useful, and some even criticized the fact that interviewers had collected too much information as most of the information needed to code a case comes from the job title, and that interviewers should focus their efforts on obtaining good information there.

Given the potential disconnect between interviewers and coders regarding what constitutes a useful occupation description, as well as the additional costs of post-interview coding, several studies have stressed the potential advantages of computer-assisted occupation coding during the interview, which would eliminate or at least minimize the need for post-interview coding (Campanelli et al. 1997; Conrad et al. 2016; Belloni et al. 2016; Helppie-McFall and Sonnega 2018). These studies cite the potential for a reduction in coding errors as the uncertainties likely to arise in a post-interview coding, due to insufficient or contradictory information provided by respondents, can be resolved by the interviewer. This, in turn, may yield a more parsimonious list of best matching occupations to choose from, and even allow the coding decision to be confirmed by the respondents themselves. If the coding instrument produces a lengthy list (or conversely, an empty list) of likely occupations to choose from, then it is much easier for the interviewer to probe for additional information during the interview than it is for any post-interview intervention to be performed. Although it is typically assumed that interviewers do not achieve the same levels of accuracy as specialist coders, with increasing experience the interviewer may develop a better idea of what constitutes a good occupational description and probe accordingly (Lyberg and Dean 1992; Campanelli et al. 1997; Conrad et al. 2016). Occupation coding during the interview is also expected to reduce costs and processing time since a smaller number of occupation descriptions will require post-interview coding. The method also has low maintenance costs as the code frame and search algorithms can be constructed and updated automatically (Hacking et al. 2006). However, a number of concerns have been raised as well, including the potential for interviewer effects and the challenge of implementing the coding instrument in a web survey, in which respondents perform the entire occupation coding process themselves without the assistance of an interviewer. Furthermore, the potential risks of mode effects when applying the coding method in a mixed-mode survey have been acknowledged elsewhere (Conrad et al. 2016; Tijdens 2014, 2015a; Tijdens and Visintin 2017).

Occupation coding of open-ended descriptions during the interview is typically implemented as a special form of computer-assisted coding in which the computer suggests the most relevant occupation code(s) to the interviewer or an online respondent. It usually follows a two-step approach. In the first step, the interviewer (or respondent) types into the open text field the job title and/or description of the occupational activity. On the basis of this verbatim text, and sometimes other input from the interview, the search engine then shows a list of best matching occupations from the code frame, from which the interviewer (or respondent) selects the most appropriate occupation. As interviewers (or respondents) enter more inputs, the search engine adapts the list to find the best matching occupations and the list of occupations becomes smaller. If the search results do not yield any likely matches, then respondents may be asked to provide further details or the case might be referred to a specialist coder post-interview. As mentioned, various matching

algorithms could be employed. The classical algorithm consults a coding index or look-up table and produces a list of appropriate categories that are identical or similar to the job title and job description information provided by the respondent using rule-based techniques (Hacking et al. 2006; Elias et al. 2014; Tijdens 2015; Tijdens and Visintin 2017; Brugiavini et al. 2017; Belloni et al. 2016; Schierholz et al. 2018; Gweon et al. 2017; Schierholz and Schonlau 2020). A more sophisticated approach uses machine-learning algorithms to identify possible occupation codes from previously coded data, known as training data (Schierholz et al. 2018; Gweon et al. 2017; Schierholz and Schonlau 2020). Benefits of combining algorithms that rely on job titles from a coding index with statistical learning algorithms trained on data from previous surveys have also been documented (Schierholz and Schonlau 2020).

Assessments of occupation coding of open-ended questions during the interview in probability-based sample surveys have been mostly positive. For example, around 80% of all occupations collected in the sixth wave of the Survey of Health Aging and Retirement in Europe (SHARE), conducted in 2015, were coded during face-to-face interviews (Brugiavini et al. 2017). A coding rate of 72% was observed by Schierholz et al. (2018) in a telephone survey in Germany in 2014, commissioned by the Institute of Employment Research (IAB). A field experiment by Statistics Netherlands in September 2003 with the Dutch Labour Force Survey (LFS) and the actual LFS in January 2004, administered face-to-face, achieved interview coding rates of 79% and 75%, respectively (Hacking et al. 2006).

Schierholz et al. (2018) also evaluated data quality by comparing telephone interview coding with office coding performed by two independent professional coders. This resulted in a high level of agreement between the office coding and the interview-coded data, which provided some reassurance that interview coding yields comparable data quality to manual coding. However, higher disagreement rates between office coders were observed for more complex occupation descriptions, to which an occupation code could not be assigned during the interview and the occupation description had to be manually coded. The authors suggest that 'simpler' occupations are more easily codable during the interview, and that more difficult descriptions are more appropriate for office coding.

However, in terms of the time needed to code occupation, concerns have been expressed that occupation coding during the interview may significantly extend interview time. Hacking et al. (2006) reported that the average duration of coding occupation using the look-up table method was 47 seconds compared to 36 seconds when only the open text information (to be coded post-interview) was collected. Schierholz et al. (2018) noted that, for respondents whose occupations were coded successfully during the interview, the duration of the interview was shortened by a few seconds compared to those who did not select one of the suggested categories and were presented with an additional follow-up question. Tijdens (2016) reported a mean time of 48 seconds to code one's occupation, using a semantic matching tool in the Wageindicator web survey, a non-probability survey on work and wages, which was a few seconds longer compared to a search tree navigation also offered on the web platform.

The self-coding of occupation based on open-text descriptions in an online setting is still rare. Insights on its feasibility have been documented for the aforementioned Wageindicator web survey, in which semantic matching, using look-up tables, has been

used to code respondents' occupations since 2015 (Tijdens 2015a). Respondents type their occupation and word-matches in the look-up table are instantly shown to them to select the most relevant match. The semantic matching tool has been noted as being preferred by respondents over search tree navigation offered on the platform alongside. Furthermore, it has been noted as being most suitable for self-coding of occupations provided that the occupational look-up database is sufficiently large (Tijdens 2015b).

3. Research Gaps and Study Questions

Given that the literature on interview-based coding of open-ended occupation descriptions comes primarily from interviewer-administered settings, mainly telephone and face-to-face surveys, the feasibility of coding occupations in probability-based online and mixed-mode surveys involving both self- and interviewer-administered modes remains unclear. To our knowledge some NSOs perform occupation coding for their own labour force surveys online during the interview (e.g., Statistics Denmark and Netherlands), however, evidence on their performance is not readily available.

In web surveys, occupation coding during the interview is challenging as there is no interviewer to provide assistance or probe for additional information. Thus, an easy-to-use interface is needed to facilitate respondent self-coding of occupation. In mixed-mode survey designs, this becomes even more challenging as the interface should be standardized across modes, so that the measurement of occupation is comparable across all respondents.

In this study, we investigate the feasibility of coding open-ended occupation descriptions during the interview by implementing a computer-assisted look-up system in a web-first sequential mixed-mode survey, in which web respondents were asked to self-code their occupation and interviewers in the telephone and face-to-face follow-up modes identified and recorded the relevant occupation code. Following the entry of key words, the coding tool searched for relevant job titles in the Standard Occupational Classification 2010 (SOC2010) coding index and offered a list of corresponding codes. If the look-up method was not successful in identifying a relevant occupation code, then the traditional post-interview occupation coding procedure using an open-ended question to describe the respondent's job tasks was employed.

In addition to the take-up rate of the look-up system, we examine indicators of its performance, such as the time to code occupation during the interview, the specificity of the allocated look-up codes, and the length of the occupational description (only asked as a follow-up question when an occupation code could not be assigned using the look-up system), which are thought to impact the quality of the collected occupation data. This study does not directly assess the quality of the collected occupation information, typically measured by reliability (the extent to which the same occupation code will be repeatedly assigned to the same case) or validity (the accuracy of the assigned code), which is a limitation to be considered in future work. Finally, we assess the extent to which respondents and interviewers influence the use of the new coding method. To our knowledge, no other study has previously investigated these issues in a web-first sequential mixed-mode survey.

Using data from the Next Steps Age 25 (wave 8) cohort study in the UK, we address the following research questions:

1. To what extent do web respondents use the look-up method to self-code their occupation, and how does this compare to interviewer administration of the look-up method in the telephone and face-to-face follow-up modes? Does the rate of self-coding in the web mode vary by device type (PC, laptop, tablet)?

Here, we expect that respondents interviewed via web will use the look-up system at a lower rate than those interviewed by the telephone or face-to-face follow-up modes. The absence of an interviewer is a key disadvantage in this case, as there is no one to motivate the respondent to engage with the coding instrument and provide guidance and/or probe for relevant details to make the look-up task more manageable. Interviewers undergo specific training with the coding system and likely have relevant experience in collecting occupation information; thus, they are more likely to be aware of what constitutes a valid occupation description than respondents. Without the assistance of an interviewer, the look-up task may become more burdensome for respondents, who may expend less effort than an interviewer would to select an occupation code, especially if the occupation is difficult to code.

The burden of self-coding is expected to be correlated with device type and, specifically, the presentation size of the look-up interface. Larger screens (e.g., desktop PCs) are likely to better handle longer look-up lists, improve visibility and limit the amount of burdensome scrolling necessary to identify the most relevant occupation, compared to smaller screens (e.g., laptop and tablet). Thus, we expect the coding rate of the look-up method to be proportional to the relative screen size and, thus, higher with desktops followed by laptops and tablets. We do not assess the use of the look-up method for smartphones, as smartphone participation was strongly discouraged in Next Steps and very few of such cases occurred.

2. Is the performance of the look-up method and characteristics of open-ended occupational descriptions – which are both linked to occupational data quality in previous research – comparable between the three sequential modes (web, telephone, and face-to-face)?

The absence of an interviewer to motivate respondents and probe for relevant details could compromise the quality of occupational coding in web surveys (Conrad et al. 2016). For example, given the cognitively demanding and time-consuming task of self-coding (Tijdens and Visintin 2017), it has been suggested that respondents (or interviewers) will select more generic occupation titles presented in classification lists, which appear to be correct but are suboptimal, rather than exert the necessary effort to choose a more specific occupation code (Schierholz et al. 2018). We thus expect a higher prevalence of generic codes for web respondents compared to respondents who are guided in the interviewer-administered modes. Generic codes refer to suboptimal descriptions or descriptions that are too abstract to be assigned to a more specific category. Such codes have a last digit of '0' or '9' at the 4-digit (unit group) level of the SOC. A last digit of '0' indicates that the 4-digit (unit) group is equivalent to the broader 3-digit (minor) group, or that there is only one unit group within the minor group and the coding could not be achieved at a more detailed level. A last digit of '9' indicates occupations "not elsewhere classified – n.e.c.", thus containing a mix of occupations which are not in sufficient numbers to merit their own unit group.

Furthermore, we expect shorter occupational descriptions to the standard open-ended question (only used as a follow-up if coding was not successful using the look-up method) by respondents on the web, compared to the interviewer-administered modes, and that the look-up procedure and the standard open-ended question will each take longer to administer for web respondents compared to respondents assisted by a trained interviewer in the telephone and face-to-face modes.

3. Do respondent attributes (e.g., sex, ethnicity, education, cohabitation status) and interviewer characteristics (e.g., sex, age, years of interviewing experience) influence the performance of the look-up method during the interview?

Based on the literature, we expect that study members' and interviewers' characteristics will affect whether the look-up method is successful in assigning an occupation code during the interview, or if post-interview coding is required. For example, Belloni et al. (2016) showed that coding errors (namely coding disagreement between manually coded verbatim responses on current and last occupation and codes assigned with the CASCOT software, while the automated coding was taken as a benchmark) were more common for male than for female respondents. For coding 'last job', errors were more likely to occur for the most educated individuals and for the self-employed. Cognitive abilities were found to play an important role in explaining coding errors for 'current job'.

It has been suggested in previous research that coder experience might affect agreement and thus coder information should be included in the analysis of coding quality (Conrad et al. 2016). Schierholz et al. (2018) also acknowledged the potential for interviewer effects on the selection of an occupation code and analyzed the extent to which interviewers correctly applied standardized interviewing techniques that were prescribed for coding occupation using behavioural coding. The authors found that many interviewers did not closely follow the rules for standardized interviews and it was rather an exception that the interviewer read out loud the exact question text and all answer options, including the last option for 'other occupation'. However, the authors noted that when the script was not followed, it was often because the interviewer already had a good understanding of the respondent's job and thus good reasons for departures from it. Nevertheless, the interplay between the interviewer and respondent was acknowledged as an important issue for future improvements of the occupation coding instrument. We therefore consider the inclusion of interviewer covariates as a strength of the present study. However, the direction of their expected impact is less clear.

4. Methods and Data

4.1. Next Steps and the Age 25 Survey

Next Steps follows the lives of 16,000 people in England born in 1989/90, sampled from state and independent schools. The sample design considered schools the primary sampling unit, with deprived schools being over-sampled by 50%. A total of 647 state and independent secondary schools as well as pupil referral units participated in the study out of 892 selected schools. Pupils from minority ethnic groups (Indian, Pakistani, Bangladeshi, Black African, Black Caribbean, and Mixed) were over-sampled to provide

sufficient base sizes for analysis. The school and pupil selection approach ensured that, within a deprivation band and ethnic group, pupils had an equal probability of selection (Department for Education 2011).

The study began in 2004, when the cohort members were aged 14. They were surveyed annually until 2010 (waves 1–7), and then in 2015–2016 when they were aged 25 (wave 8). The interviews for the first four waves were conducted face-to-face, and from wave 5 onwards a sequential mixed mode approach – online, followed by telephone, and then face-to-face interviews – was used. Next Steps has collected information about cohort members' education and employment, economic circumstances, family life, physical and emotional health and wellbeing, social participation and attitudes. A total of 15,531 cohort members were issued to field in the most recent age 25 survey, achieving a response rate of 51% with 7,707 completed interviews (4,797 online, 690 telephone, and 2,220 face-to-face) (Centre for Longitudinal Studies 2017).

4.2. Occupation Coding System

Economic activity data has been collected in the study since its initial wave at age 14 and occupation data in particular since study members were aged 16 – that is, when they reached the compulsory school leaving age and were eligible to start an apprenticeship or traineeship, or spend 20 or more hours per week working or volunteering while in part-time education or training. Occupation has since been captured with open-ended questions asking about the title of their job and a description of what they mainly do in their job. This information has been subsequently coded post-interview by professional office coders.

Labor market entry is a key milestone in the cohort's transition from adolescence to young adulthood, and thus work participation was a key theme in the age 25 survey. In this survey, occupation was captured by a text-based search and coding method during the interview. Following an open-ended question about their job title ('What is your current job title?'), respondents in the web survey were asked to enter key words into a search box describing what they mainly do in their job, then select the most appropriate response option from a list of occupations generated by the search system (Figure 1). The input string of words was matched against a concatenated string of the fields describing the job

Fig. 1.   Screenshot of the look-up question in the web survey.

title in the Standard Occupational Classification 2010 (SOC2010) coding index (containing over 27,000 job titles) as a look-up. This concatenated string included the indexing word (usually the word describing the core set of tasks that characterize a job), occupational, industrial, and additional qualifying terms. The occupational qualifying term is separated from the indexing word by a comma (e.g. teacher, head). Industrial qualifying terms are shown within brackets and can take the form of an industry or branch of industry in which the occupation lies (e.g. teacher, head, (secondary school)). Additional qualifying terms usually indicate the type of material worked with, the machinery used or the processes involved, or can take the form of professional qualifications (Office for National Statistics 2010b). As the indexing word is rarely enough to enable the job title to be correctly coded, the additional qualifying terms aimed to make the search more specific, and, in turn, the coding more accurate. The look-up method used a "word-chunk" search system processing the input string as a string of chunks (comprised of at least three characters) and searching for each chunk in the SOC2010 job title index (i.e., a simple lookup of the job titles containing all of the word chunks). There was no pre-processing (standardization) of the key words entered in preparation for matching, and no amendments were made to the underlying job title list.

Following the entry of the key words, occupation codes at the most detailed 4-digit unit group level of the classification were displayed in alphabetical order. As the SOC2010 index of job titles includes occupation codes with a '0'-ending digit (indicating a single unit group within the minor group) and a '9'-ending digit (indicating 'not else classified'), these respective job titles were displayed if relevant to the search. The procedure was similar in the interviewer-administered telephone and face-to-face modes, except interviewers entered the key words into the search box, read out the list of occupation search results to the respondent, and selected the most appropriate occupation from the list.

A further instruction, aimed at handling long lists of occupations, stated that using extra words or parts of words will narrow down the list of displayed options. For example, just typing in "teacher" would bring up a long list of possible occupations, but entering additional search terms, for example, "teacher secondary", would narrow down the list of options (Figure 2).
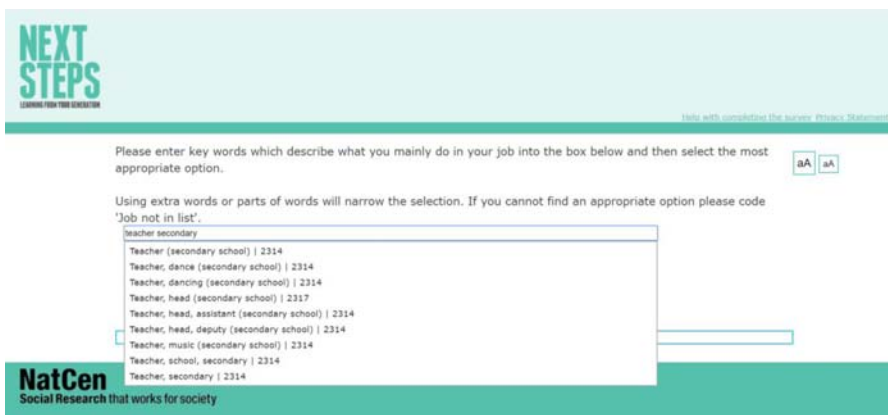

Fig. 2.   *Screenshot of the look-up question and search results for "teacher secondary" in the web survey.*

If an appropriate option was selected during the interview, the associated code from the SOC2010 coding index was automatically assigned, replacing the respondent's entry of key words (required to perform the search). After trying different search terms, if an appropriate option was not found (e.g., no search results were presented, or none of the presented search results were considered appropriate by the respondent or interviewer), then respondents (or interviewers) were asked to select the 'Job not in the list' option and answer a follow-up standard open-ended question describing their job ("Please describe in your own words what you mainly do in your job."). For this question, respondents (or interviewers) were encouraged to provide (or probe for) full details (for example, the type of work) to allow office coders to accurately identify the correct SOC code after the interview.

There was no difference in the design or presentation of the look-up question or standard open-ended job description question between the self- and interviewer-administered modes other than a minimal altering of the question wording and accompanying instructions, so that they were appropriate for either the respondent or interviewer. Telephone and face-to-face interviewer scripts included an additional instruction 'to probe as required'.

4.4.   *Statistical Analysis*

To address the first research question, we report percentages of respondents who were assigned an occupation code during the interview using the look-up coding method and the complementary set that were referred to traditional office coding post-interview. Using chi-squared tests, we show how the look-up rates varied by the sequentially offered modes (online, telephone and face-to-face). We also use chi-squared tests to evaluate differences in the coding rates by self-reported device type: desktop computer, laptop and tablet in the web interviews. A very small number of web respondents ($n = 25$) reported completing the survey on a smartphone, even though this practice was discouraged. These few cases are excluded from all analyses.

To address the second research question, we perform three separate analyses, which we compare across the three modes. First, using chi-squared tests we assess the difference in the prevalence of (suboptimal) generic codes (i.e., codes with a last digit of '0' or '9' at the 4-digit level) compared to more specific occupation codes collected using the look-up method. The second analysis evaluates the length of the open-text descriptions (measured by the total number of characters) provided by respondents to the third open-ended question asked only if an occupation code was not assigned during the interview using the look-up method. We use a Wald test to evaluate differences in the mean description length between the sequentially offered modes. The third analysis assesses the time that was required to code occupation using the look-up method and provide an open-text description in the case of inability to assign an occupation code during the interview. Using a Wald test, we evaluate if the mean time (measured in seconds) differed across the web, telephone, and face-to-face modes.

To address the third research question, we first fit a multivariable logistic regression model on whether or not an occupation code could be assigned during the interview (using the look-up method). We report crude (unadjusted for other characteristics) and adjusted (for all characteristics) odds ratios (OR) and 95% confidence intervals (CI) for the effects of the following respondent characteristics: sex, ethnicity (white/non-white), if ever attended

university and if in a cohabiting relationship by age 25, and survey mode. We further include interactions between each of the respondent characteristics and response mode.

To more fully account for mode-specific effects and differing levels of information available for each mode, including interviewer characteristics in the face-to-face mode, we fit separate logistic regression models on the assignment of an occupation code during the interview for each of the three response modes. A two-level random intercepts model is used for the face-to-face respondents to account for respondents nested within interviewers and to assess the effects of the following interviewer characteristics: sex, age (recoded: 49 years or younger, 50–59, 60–69, 70+ years) and years of experience (recoded: 1 year or less, 2–5 years, 6–9 years, 10 or more years). The intraclass correlation coefficient (ICC) is also reported for the face-to-face model as an approximate measure of the "interviewer effect" or the proportion of variation in the coding outcome attributable to the interviewer level. This analysis is only possible for the face-to-face interviews as interviewer IDs were not recorded for telephone interviews.

We note that Next Steps did not randomly assign cohort members to interviewers. Thus, the reported ICC may reflect both area and interviewer effects which are confounded. The lack of an interpenetrated design is a limitation which we attempt to address by including respondent and interviewer characteristics in the model. Still, it is plausible that further controls are needed to isolate the pure interviewer effect. Thus, we interpret the ICC with caution.

All analyses addressing the three research questions account for the complex sample design of the Next Steps study using the SVY commands in Stata 16.0 and control for selection into each sequential mode using weights, which we describe next. Descriptive statistics for all variables used in the analysis are supplied in online supplemental material Table S1.

4.5. Accounting for Sequential Mode Selection

To evaluate and compare the performance of the look-up method across the three survey modes, it is useful to control for differential nonresponse at each stage of the sequential mixed-mode design. Several methods have been proposed to control for differential nonresponse in mixed-mode surveys (Vannieuwenhuyze and Loosveldt 2013; Vannieuwenhuyze et al. 2014; Klausch et al. 2017). One of the most common methods is to use selection weights that are based on the (estimated) propensity of a respondent to participate in each mode (Hox et al. 2015). The mode propensities are typically estimated from a generalized linear model (e.g., probit, logit), with each mode treated as a possible outcome, conditioning on available covariate information which, in longitudinal studies, is often limited to fixed baseline characteristics (e.g., demographics) and/or data collected from the previous wave.

We adopt and extend this method by applying a data-driven nonresponse weighting procedure. Instead of using only baseline covariates or covariates collected only in the previous wave of Next Steps, we use covariates selected from all (seven) previous waves to adjust for mode selection in the Next Steps age 25 (wave 8) survey. A two-step approach was implemented separately and sequentially for each of the three response modes, starting with web followed by telephone and then face-to-face. In the first step, seven

multivariable log-binomial regressions predicting nonresponse at wave 8 were fitted, each regression containing only predictor variables collected from one of the seven prior waves of Next Steps. All statistically significant ($p < 0.05$) predictors were retained for the second step of the procedure.

In the second step, all of the variables retained from the first step were imputed to produce a complete dataset of predictors. These wave-specific predictors then entered into a series of log-binomial regression models predicting nonresponse at wave 8, each model building on the previous one by incorporating additional variables from the subsequent wave. For example, the first model of nonresponse at wave 8 included only predictors from wave 1, then wave 2 predictors were added into the next model, and so on. After introducing a given set of wave-specific predictors, these predictors were checked for statistical significance ($p < 0.05$). If a current-wave predictor was no longer statistically significant after controlling for the predictors from past waves, it was dropped from the model. Only predictors which remained significant after controlling for predictors from the past waves were retained. This was done to maintain the temporal sequence of the predictors available in the longitudinal data.

All retained variables (shown in online supplemental material Tables S2–S4) were then used to create propensity score adjustment weights for mode-specific unit nonresponse. The propensity to respond at each stage of the sequential mixed-mode design (web, telephone, and face-to-face) was calculated separately for each sample unit. The estimated propensity scores were then sorted into quintiles. The nonresponse adjustment weight was then calculated as the inverse of the average propensity score in each quintile. This process was performed separately for each response mode in their sequential order, yielding three nonresponse adjustment weights. The final analysis weights were then computed as the product of the Next Steps base weight and the nonresponse adjustment weights generated from the above procedure.

5. Results

5.1. *Look-Up Method and Office Coding Rates*

The occupation coding rates are presented in Table 1. First, we report the percentage of respondents who were assigned an occupation code during the interview using the look-up coding method. Overall, across the three modes, 82.0% of respondents were successfully assigned an occupation code using the look-up method. The remaining respondents who were not assigned an occupation code during the interview were either assigned one by an office

Table 1. Occupation coding rates by survey mode.

Occupation coding method	Web % (n)	Telephone % (n)	Face-to-face % (n)	Total % (n)
Look-up coding	90.3 (3,580)	90.6 (493)	69.2 (1,180)	82.0 (5,253)
'Successful' office coding	8.7 (356)	8.1 (50)	30.5 (483)	17.4 (889)
'Unsuccessful' office coding	0.3 (9)	0.8 (3)	0.0 (1)	0.2 (13)
Refused	0.7 (30)	0.5 (3)	0.2 (7)	0.5 (40)
Total	100 (3,975)	100 (549)	100 (1,671)	100 (6,195)

Notes: Occupation coding outcome vs. survey mode: $\chi_6^2 = 508.8212$, $p < 0.000$. Results are weighted to account for selection into each of the respective sequentially administered modes.

coder after the interview (17.4%), or lacked sufficient information to be coded due to failure to answer the occupation questions (0.5%) or inability of the office coder to identify an appropriate code (0.2%).

The percentage of respondents who successfully used the look-up method varied significantly by mode (Table 1). Contrary to our expectation, look-up rates in the web mode were among the highest with 90.3% of respondents able to self-code their occupation during the interview. This was comparable to the telephone mode, in which 90.6% of respondents were coded during the interview (web versus telephone: $\chi_3^2 = 5.1816$, p = 0.3679). The look-up rate was significantly lower among face-to-face respondents, who were assigned an occupation code in only 69.2% of interviews (web vs. face-to-face: $\chi_3^2 = 450.4548$, p < 0.000; telephone versus face-to-face: $\chi_3^2 = 82.3264$, p < 0.000).

The assignment of an occupation code with the look-up method or with standard office coding varied slightly by (major 1-digit) occupation groups. Namely, occupations in Major Group 8 (Process, plant and machine operatives) appeared to be coded at a higher rate post-interview than during the interview, whereas occupations in Major Group 2 (Professional occupations) were more often coded using the look-up method (results not shown). All other occupation groups were assigned at similar rates by coding method.

Coding rates by device type are reported for the web respondents in Table 2. Look-up coding rates were slightly higher for desktop computers (93.1%) compared to devices with likely smaller screens, including laptops (92.1%) and tablets (90.9%). This pattern followed our expectations; however, the overall differences are not statistically significant ($\chi_2^2 = 3.06$, p = 0.541).

5.2. Other Performance Measures

Next, we assess several other indicators of performance for the look-up method and the standard occupation question (used in post-interview coding) likely to impact the collected occupation data. First, we examine the prevalence of generic vs. specific codes, then we look at the length of open-text descriptions to the open-ended occupation question (if the look-up method was unsuccessful). Lastly, we evaluate the mean time it took respondents and interviewers to select an appropriate occupation using the look-up method and answering the open-ended occupation question (if applicable).

5.2.1. Specific Versus Generic Codes

The allocation of an occupation code does not necessarily mean that an optimal code has been found, since some occupation codes refer to more general or abstract occupations

Table 2. *Occupation coding rates by self-reported device type in web interviews.*

Coding method	Desktop % (n)	Laptop % (n)	Tablet % (n)	Overall % (n)
Look-up coding	93.1 (691)	92.1 (1,818)	90.9 (954)	91.9 (3,463)
Office coding	6.9 (54)	7.9 (156)	9.1 (116)	8.1 (326)
Overall	100 (745)	100 (1,974)	100 (1,070)	100 (3,789)

Notes: Refusals (n = 30) and non-coded post-interview cases (n = 9) are excluded. Cases with missing information on device type (n = 4) and cases completed on a smartphone or other device (n = 25) are also excluded. A further 118 cases assigned a non-applicable code for self-reported device type were also excluded. Results are weighted to account for selection into the initially administered web mode.

compared to those with more detailed or specific occupation meanings. If a general occupation code is selected, when a specific code exists, this may result in misclassification and contribute to bias in the analysis of occupation data. Of course, there are legitimate reasons why a general occupation code may be chosen rather than a more specific one, not driven by lack of respondent engagement; for example, if a specific code doesn't exist in the SOC for a person's occupation, or if the occupation is not prevalent enough in the population to warrant its own category. However, assessing the validity of the assigned codes is beyond the scope of this analysis and here we focus on the prevalence of specific and generic occupation codes across response modes and coding methods.

Percentages of respondents assigned specific and generic (4-digit) occupation codes during the interview (using the look-up method) and via post-interview office coding, by response mode, are presented in Table 3. Overall, we find no statistically significant differences in the assignment of specific or generic occupation codes by response mode for either look-up coding ($\chi^2_4 = 5.63$, p = 0.571) or office coding ($\chi^2_4 = 6.50$, p = 0.157). About 85% of web respondents selected a 'specific' code (not a '0' or '9' last digit) using the look-up method. Contrary to our expectation, this was comparable to the percentages in the telephone (82.8%) and face-to-face (84.6%) modes. About 84.2% of web respondents referred to post-interview office coding were assigned a specific code, which was also comparable to the telephone (88.9%) and face-to-face (79.5%) respondents.

5.2.2. Length of Open-Text Occupation Descriptions for Post-Interview Coding

Previous research on the detail of the information collected via standard open-ended occupation questions and their usefulness for successful office coding is mixed. For example, longer descriptions were found by Massing et al. (2019), Helppie-Mcfall and Sonnega (2018), Conrad et al. (2016), and Bergmann and Joye (2005) to be less reliably coded than shorter ones, and the additional information (either in longer descriptions or through additional questions or probes) was found to be associated with lower levels of coder agreement. However, it was also observed that this effect was stronger for particular occupation terms (Conrad et al. 2016). Belloni et al. (2016), on the other hand, highlighted the importance of auxiliary information, which substantially increased the level of detail (i.e., number of digits) at which the observations were coded. Similarly, Campanelli et al. (1997) noted that the combination of both the job title and job description used in automated coding led to results that were comparable to manual coding, which was used as a benchmark.

Our analysis does not provide evidence about the accuracy or reliability of coding based on the length of the occupation description collected from respondents who were not assigned an occupation code during the interview, other than the fact that almost all descriptions collected with the standard open-text question, irrespective of mode, were successfully coded post-interview. Rather, our analysis compares the length of the occupation descriptions provided in the self-administered web mode to the two interviewer-administered modes. It has been previously hypothesized that obtaining detailed occupation descriptions from respondents is more challenging in passive self-administered modes because there is no interviewer to probe for more information. Thus, there is a higher risk that occupation descriptions supplied in self-administered modes will be shorter and less sufficient in terms of detail (Conrad et al. 2016).

Table 3. *Percentages and 95% confidence intervals (CI) of specific and generic occupation (4-digit) codes by coding method and survey mode.*

Occupation code	Coding method				Office coded % (95% CI)			
	Look-up % (95% CI)							
	Overall	Web	Telephone	Face-to-face	Overall	Web	Telephone	Face-to-face
0-ending	5.3 (4.4-6.3)	4.7 (3.8-5.9)	6.2 (3.9-9.9)	6.0 (4.4-8.1)	5.7 (3.9-8.5)	3.3 (1.9-5.6)	0.4 (0.1-3.0)	6.9 (4.4-10.9)
9-ending	10.0 (9.0-11.1)	10.3 (9.0-11.6)	11.0 (7.4-16.2)	9.4 (7.6-11.5)	13.2 (10.3-16.6)	12.5 (0.8-19.2)	10.6 (4.4-23.5)	13.53 (10.1-17.9)
Specific (not ending with '0' or '9')	84.7 (83.3-86.0)	85.0 (83.4-86.6)	82.8 (76.9-87.5)	84.6 (82.0-86.9)	81.1 (77.1-84.5)	84.2 (77.7-89.2)	88.9 (76.2-95.3)	79.5 (74.3-83.9)

Note: Results are weighted to account for selection into each of the respective sequentially administered modes.

On the contrary, we find that in fact web respondents provided, on average, longer descriptions (62 characters) to the open-text occupational question than their telephone and face-to-face counterparts (both 45 characters). The differences between the web mode and the interviewer-administered modes were statistically significant at the 0.05 level (web vs. telephone: $F_{1,397} = 3.92$, p = 0.048; web vs. face-to-face: $F_{1,397} = 10.11$, p = 0.002). No significant difference was found between the telephone and face-to-face modes ($F_{1,397} = 0.00$, p = 0.981).

5.2.3.   Time to Select An Occupation or Provide An Occupation Description

Although the primary aim of the look-up method is to increase the accuracy and cost efficiency of occupation coding, it has been emphasized in previous research that, if interview coding is to replace office coding, then it is important that the procedure does not significantly increase interview duration, as longer interviews are more expensive and burdensome for participants (Schierholz et al. 2018). Moreover, gains in cost efficiency due to the reduction of manual coding may be partly or completely offset if interview duration increases significantly (Hacking et al. 2006).

Overall, across the three modes, the look-up method used in Next Steps took, on average, 43 seconds regardless of whether an occupation category was chosen. This varied significantly by mode and took longer in web (47 seconds) compared to telephone (35 seconds) and face-to-face (34 seconds) (web vs. telephone: $F_{1,616} = 42.10$, p < 0.000; web vs. face-to-face: $F_{1,616} = 79.81$, p < 0.000). The difference between telephone and face-to-face was not statistically significant (telephone vs. face-to-face: $F_{1,616} = 0.36$, p = 0.550).

For those respondents who could not be assigned an occupation code using the look-up method and who had to describe their occupations in the follow-up open-text question, it took on average an additional 38 seconds to answer this question. The time to write an occupation description in the open-text question also varied by mode and took longer in web (41 seconds), followed by face-to-face (35 seconds) and telephone (27 seconds). The difference between web and telephone is statistically significant ($F_{1,334} = 7.80$, p = 0.001), whereas the differences between web and face-to-face ($F_{1,334} = 3.38$, p = 0.067) and telephone and face-to-face ($F_{1,334} = 2.86$, p = 0.092) are both marginally significant.

We note that our timing analysis is performed on a reduced sample due to loss of timing information in 11% of the Next Steps face-to-face interviews (Centre for Longitudinal Studies 2017). It also excludes outliers, defined as observations above the 99th percentile.

5.3.   *Correlates of Occupation Coding During the Interview*

We next explore the extent to which study members' characteristics and interviewer attributes (observed only for the face-to-face interviews) influenced whether or not an occupation code could be assigned during the interview (via the look-up method). We look at the effects of these characteristics separately for each of the three sequential response modes (i.e., by fitting mode-specific models) to allow for the different numbers of respondent- and interviewer-level characteristics to vary across the different modes, which would otherwise be obscured in a single combined model restricted to attributes observed in all three modes. Nevertheless, a combined model with these restrictions is presented in

online supplemental material, Table S5, for the interested reader. However, only the mode-specific models are presented and interpreted below.

Table 4 shows the crude (unadjusted for other characteristics) and adjusted odds ratios reflecting the association between each of the respondents' (and, if available, interviewers') characteristics and successful use of the look-up method to assign an occupation code during the interview, presented separately for each response mode.

5.3.1. Online Model

Starting with the online model (Table 4, column A), we find that successful coding of occupation during the web interview (via the look-up method) is related to respondents' ethnic background, whether or not they attended university by age 25, and whether or not they are in a cohabiting relationship at age 25. White study members had 2.2 times higher odds of being coded during the interview compared to non-white study members (OR = 2.20, 95% CI: 1.44-3.36). Those who attended university by age 25 were 1.7 times more likely to be assigned an occupation code during the interview compared to those who have not attended university by that age (OR = 1.70, 95% CI: 1.16-2.50). And those living with a partner, at age 25, were 1.6 times more likely to be assigned an occupation code during the interview compared to those not living with a partner. There was no evidence that study members' sex or the device type used to complete the web survey were related to successful use of the look-up method, after accounting for all other characteristics of interest.

5.3.2. Telephone Model

Continuing with the telephone model (Table 4, column B), we find no strong evidence that respondents' demographic characteristics influenced the likelihood of being assigned an occupation code during the interview. Respondents' sex, ethnic background, cohabitation status and whether or not they have attended university by age 25 were not associated with receiving an occupation code during the telephone interview.

5.3.3. Face-to-Face Model

For the face-to-face model (Table 4, column C), there was also no evidence that study members' demographics influenced their likelihood of being assigned an occupation code during the face-to-face interview, apart from their cohabitation status. As in the web interviews, participants in a cohabiting relationship were more likely to receive an occupation code during the interview (OR = 1.67, 95% CI: 1.08-2.57). However, as expected, there was strong evidence that interviewers influenced whether an occupation code was assigned during the face-to-face interviews. In particular, males, older and less experienced interviewers were less likely to successfully assign an occupation code to study members during the interview using the look-up method. Compared to males, female interviewers had over three times higher odds of assigning an occupation code during the interview (OR = 3.19, 95% CI: 1.92-5.31). The odds of assigning an occupation code notably decreased with increasing age of the interviewer. Compared to interviewers with less than a year of experience in the fieldwork agency, those with experience between two and ten years were considerably more likely to successfully use the look-up method to assign an occupation code. Interviewers with over ten years of

Table 4. Odds ratios (OR) and 95% confidence intervals (CI) of successful occupation coding during the interview (via the look-up method) on respondent and interviewer characteristics, by survey mode

	(A) Web (n = 3,789)		(B) Telephone (n = 543)		(C) Face-to-face (n = 1,576)	
	Crude OR (95% CI)	Adjusted OR (95% CI)	Crude OR (95% CI)	Adjusted OR (95% CI)	Crude OR (95% CI)	Adjusted OR (95% CI)
Respondent characteristics						
Female	0.87 (0.60-1.24)	0.93 (0.65-1.34)	0.66 (0.34-1.28)	0.68 (0.36-1.30)	0.92 (0.0.69-1.23)	0.73 (0.45-1.19)
(REF: Male)						
White	2.09*** (1.43-3.08)	2.20*** (1.44-3.36)	0.83 (0.42-1.64)	0.67 (0.32-1.42)	0.89 (0.64-1.22)	0.84 (0.48-1.48)
(REF: Non-white)						
Ever attended university						
(REF: No)	1.53* (1.08-2.17)	1.70** (1.16-2.50)	0.60 (0.30-1.19)	0.61 (0.31-1.21)	0.84 (0.62-1.15)	0.67 (0.40-1.13)
Cohabiting	1.76** (1.21-2.54)	1.55* (1.06-2.25)	1.07 (0.52-2.20)	1.17 (0.56-2.41)	1.34 (1.02-1.77)	1.67* (1.08-2.57)
(REF: No)						
Laptop (REF: Desktop)	0.91 (0.66-1.26)	0.81 (0.45-1.46)	–	–	–	–
Tablet	0.65* (0.46-0.91)	0.71 (0.41-1.24)	–	–	–	–
(REF: Desktop)						
Interviewer characteristics						
Female	–	–	–	–	4.08*** (2.46-6.78)	3.19*** (1.92-5.31)
(REF: Male)						
Age: 50-59 years	–	–	–	–	0.10*** (0.04-0.29)	0.12*** (0.04-0.36)
(REF: 49 or younger)						
Age: 60-69 years	–	–	–	–	0.05*** (0.02-0.14)	0.08*** (0.03-0.24)
(REF: 49 or younger)						
Age: 70+ years	–	–	–	–	0.01*** (0.00-0.04)	0.02*** (0.01-0.07)
(REF: 49 or younger)						
Experience: 2-5 years	–	–	–	–	3.35** (1.36-8.26)	3.50* (1.32-9.30)
(REF: 1 year or less)						
Experience: 6-9 years	–	–	–	–	1.22 (0.57-2.59)	2.60* (1.15-5.87)
(REF: 1 year or less)						
Experience: 10+ years	–	–	–	–	0.25*** (0.11-0.54)	0.68 (0.31-1.50)
(REF: 1 year or less)						

Notes: *** p < 0.001, ** p < 60.01, * p < 0.05, + p < 0.1. Results are weighted to account for selection into each of the respective sequentially administered modes.

experience were less likely compared to those with less than a year of experience to code respondents during the interview, although this effect was not statistically significant (OR = 0.68, 95% CI: 0.31-1.50). Further to the above observations, the estimated ICC (rho = 0.67) showed that 67% of the variability in the look-up coding outcome, after accounting for interviewers' and participants' characteristics, was due to variability between interviewers.

6. Discussion

This study illustrated the feasibility of coding occupations during the interview using a standardized coding look-up system implemented in a large-scale sequential mixed-mode (web, telephone, face-to-face) survey of young adults in the UK. Occupation coding is particularly important in this population in the stages of transition into the labor market. To our knowledge, this is the first study to present findings on the feasibility of occupation coding during the interview in a mixed-mode survey, in which the occupation coding procedure is performed by respondents themselves in the first offered web mode, and by interviewers in the follow-up telephone and face-to-face modes. The design and implementation of occupation coding during the interview is challenging in online surveys, in the absence of an interviewer to guide or motivate respondents to perform the self-coding. However, these are even more challenging in mixed-mode studies, for which standardization of the measurement of occupation is desired to the maximum extent possible to ensure comparability of the collected data across modes.

The study yielded five main findings. First, the look-up coding method was considered highly effective as 82% of all respondents were assigned an occupation code during the interview. This is rather similar to the coding rates achieved in previous studies (Hacking et al. 2006; Brugiavini et al. 2017; Schierholz et al. 2018), acknowledging the differences in study populations, designs and occupation coding frames. This result suggests high potential for cost savings as only the remaining 18% of respondents required subsequent office coding.

Second, the success of the look-up coding method varied significantly by survey mode. It achieved a rate of about 90% in the web and telephone interviews, and about 70% in the face-to-face interviews. This finding contradicted our expectation that the look-up coding method would perform better in the interviewer-administered modes compared to the self-administered web mode. Nevertheless, it is a promising finding and suggests that respondents are not overly burdened with the task of looking up and assigning an occupation code to themselves. This is also a particularly timely finding, as web surveys are becoming more popular in survey research and established interviewer-administered surveys are increasingly transitioning to more online and mixed-mode data collection, as Next Steps has done since wave 5 onward.

Third, despite concerns that coding occupational descriptions may perform differently and sub-optimally in a web survey compared to interviewer-administered modes, we found them to be comparable across both mode types. There was no difference in the proportion of generic occupation codes (i.e., codes with last digit '0' or '9' at the 4-digit level of SOC) assigned during the interview or in post-interview office coding across the different survey modes. However, some occupations were more likely to be coded using the look-up method,

such as Professional occupations, than others, namely Process, plant and machine operatives. Almost all occupation descriptions captured with the open-text question (due to failure of the look-up method) were successfully office-coded across all modes.

Fourth, web respondents were about 11 and 13 seconds slower, on average, in using the look-up method to identify an appropriate occupation code compared to telephone and face-to-face respondents, respectively. Web respondents also took longer to describe their jobs if they could not assign themselves an occupation code using the look-up method: 16 seconds longer than telephone and seven seconds longer than face-to-face respondents. However, the longer online durations led to longer descriptions (about 17 more characters, on average) compared to those recorded in the telephone and face-to-face interviews. Again, these observations give positive insights into implementing occupation coding in self-administered surveys; namely, that self-coding of occupation does not appear to substantially extend the interview duration and respondents who are unable to self-code themselves tend to enter more details in the open-text form compared to interviewers. The difference in the timing to code occupation with the look-up method or write a detailed occupation description was expected as web respondents may require more time to read, comprehend and respond to the requests, than respondents who are assisted by a trained interviewer who is already familiar with the procedure. This, however, could also be a result of less time pressure on the web.

Lastly, we found that both study members' and face-to-face interviewers' characteristics influenced whether an occupation code was assigned during the interview, even after accounting for selection into each sequential mode. Study members' ethnic background, university participation and cohabitation status affected occupation coding during the online interviews; while there was no evidence that these characteristics, apart from cohabitation status, influenced occupation coding during the telephone and face-to-face interviews. We observed a notable interviewer effect in the face-to-face interviews: interviewers' sex, age and years of interviewing experience strongly impacted the likelihood of assigning an occupation code during the interview, as males, older and less experienced interviewers were less likely to succeed in assigning a code using the look-up method.

The effect of interviewers in the face-to-face survey is concerning and raises the question of whether the look-up coding method used during the interview may be more burdensome for interviewers to administer to respondents than for respondents to administer to themselves in a self-administered setting. However, use of the look-up method did not appear to be as problematic for telephone interviewers who performed their interviews from a centralized telephone unit under continuous monitoring by supervisors, suggesting that the higher coding rates in the telephone interviews may have been influenced by the tighter level of control and supervision of the interviewers. Although unstandardized interviewer behavior does not necessarily have a negative impact on data quality (Schierholz et al. 2018), for those interviewers for whom the look-up method was less successful, further training and supervision may be needed. The fact that all interviewer-collected descriptions were successfully coded post-interview raises the question of whether these interviewers invested the necessary effort in using the look-up method to assign an occupation code, as opposed to reverting to their prior experience and habits of collecting occupation information using standard open-ended questions only.

In our view, this suggests that monitoring of interviewer performance and more specialized training on the use of within-interview coding methods is needed to improve their successful application. The training should make clear to interviewers the benefits of using the new coding system (e.g., processing time reduction, respondent confirmation of assigned code, etc.), while monitoring should be used as part of a feedback loop to continuously improve the application of the coding system during the field period.

Our findings support the existing research that occupation coding during the interview reduces the need for (and associated costs of) post-interview office coding. This is because the large majority of occupations could be coded using the look-up method in each of the three survey modes. However, this finding should be weighed against the added costs of increasing the length of the interview, which could be valued differently depending on the mode of administration and other survey constraints. Furthermore, when the look-up method was not successful, almost all occupation descriptions collected via the open-text question were successfully coded post-interview, which is indicative of the quality of the verbatim information provided by respondents, including that which the web respondents provided without interviewer assistance. This is encouraging as previous research has noted that "the largest source of error lies in shortcomings of the verbatim raw material," as opposed to errors resulting from coding (Hoffmann et al. 1995, 13). It is also positive that web survey respondents provided lengthier descriptions, on average, than the descriptions recorded by interviewers in telephone and face-to-face modes, which is indicative of respondents' engagement.

It has been hypothesized by Conrad et al. (2016) that, as writing or typing requires more effort than speaking for most people, it could be the case that occupation descriptions might be shorter in self-administered (visual) modes, and flagged this as an area – especially with the growth of online surveys – that warrants further study. The fact that we find the opposite effect – that web respondents offer longer descriptions – is reassuring, particularly for the more complex occupations which respondents are unable to locate using the look-up method. As there is no interviewer to probe for more specific information, providing longer descriptions during the interview to facilitate post-interview coding may be more useful than shorter descriptions. However, acknowledging that it may not necessarily be the length of the description that leads to an optimal occupation code (Conrad et al. 2016), survey designers may benefit from offering more specific instructions (including examples) to respondents and interviewers about what constitutes a good occupational description. In addition to more specific instructions, survey designers may consider following up on suboptimal or generic coding (i.e., allocation of a code ending '0' and '9') during the interview with an open-text question for more details. This could potentially enable allocation of a more specific occupation code post-interview with the generic code used as a starting point.

This research complements the existing literature with evidence about the feasibility and effectiveness of occupation coding during the interview in a large-scale, probability-based online and mixed-mode survey. It also provides insights on the performance of coding during the interview and the characteristics of the provided occupation descriptions which are likely to impact occupation data quality. To our knowledge, such an assessment has not surfaced in the research literature. Our work also identifies respondent and interviewer factors that affect the performance of the coding method during the interview and suggests ways for improvement.

There are, however, limitations that could be addressed in future work. Next Steps uses a sequential mixed-mode design, which makes it difficult to remove selection mode effects from measurement mode effects. We addressed this limitation by performing an extensive data-driven "back-door" weighting procedure utilizing seven waves of Next Steps data and numerous covariates to adjust for selection into each phase of the sequential mode design. However, there is still the potential that some factors influencing selection into mode were unaccounted for in the weighting procedure. Another limitation is that this study was performed on a panel population in its eighth wave of data collection. This population is likely to be more cooperative and perhaps more patient in engaging with the occupation coding system, than a freshly recruited sample of the general population. Nonetheless, it is reassuring that the look-up rates observed here were comparable to those observed in other studies (Hacking et al. 2006; Brugiavini et al. 2017; Schierholz et al. 2018). Furthermore, the study lacked relevant pieces of information that would provide further insights on the application of the coding method, including characteristics of the telephone interviewers, the length and content of the terms entered into the look-up search box, as well as change-logs to view the iterative process that respondents and interviewers undertook to identify an appropriate occupation code. Lastly, this study did not directly assess the quality (e.g., validity, reliability) of the occupation codes assigned using the look-up method, as the allocation of an occupation code does not necessarily imply that the optimal code was assigned. We plan to address these issues in future rounds of Next Steps and encourage future studies and other survey institutions to consider them as well. Future work is also needed on developing a theoretical framework for occupation coding.

7. References

Belloni, M., A. Brugiavini, E. Meschi, and K. Tijdens. 2016. "Measuring and detecting errors in occupational coding: an analysis of SHARE data." *Journal of Official Statistics*, 32(4): 917–945. DOI: https://doi.org/10.1515/jos-2016-0049.

Bergmann, M.M., and D. Joye. 2005. "Comparing Social Stratification Schemata: CAMSIS, CSP-CH, Goldthorpe, ISCO-88, Treiman, and Wright." *Cambridge Studies in Social Research* 10: 1–35. Available at: https://www.sociology.cam.ac.uk/system/-files/documents/cs10.pdf (accessed October 2019).

Brugiavini, A., M. Belloni, R.E. Buia, and M. Martens. 2017. *The "Job Coder"*. In SHARE Wave 6: Panel innovations and collecting Dried Blood Spots. Edited by F. Malter and A. Börsch-Supan. Munich: MEA, Max Planck Institute for Social Law and Social Policy: 51–70. Available at: http://www.share-project.org/uploads/tx_sharepublications/ 201804_SHARE-WAVE-6_MFRB.pdf (accessed October 2019).

Burstyn, I., A. Slutsky, D.G. Lee, A.B. Singer, Y. An, and Y.L. Michael. 2014. "Beyond Crosswalks: Reliability of Exposure Assessment Following Automated Coding of FreeText Job Descriptions for Occupational Epidemiology." *The Annals of Occupational Hygiene* 58(4): 482–492. DOI: https://doi.org/10.1093/annhyg/meu006.

Campanelli, P., K. Thompson, N. Moon, and T. Staples. 1997. "The Quality of Occupational Coding in the United Kingdom." In *Survey Measurement and Process Quality*. Edited by L. Lyberg, P. Biemer, M. Collins, E. De Leeuw, C. Dippo, N. Schwarz, and D. Trewin: 437–453. New York: Wiley.

Cantor, D., and J.L. Esposito. 1992. "Evaluating Interviewer Style for Collecting Industry and Occupation Information." In Proceedings of the Section on Survey Methods, American Statistical Association: 661–666. Available at: https://www.bls.gov/osmr/research-papers/1992/pdf/cp920010.pdf (accessed March 2021).

Centre for Longitudinal Studies. 2017. *Next Steps Age 25 Survey*. Technical Report. University College London. Available at: http://doc.ukdataservice.ac.uk/doc/5545/mrdoc/pdf/5545age_25_technical_report.pdf (accessed October 2019).

Conrad, F., M. Couper, and J.W. Sakshaug. 2016. "Classifying Open-Ended Reports: Factors Affecting the Reliability of Occupation Codes." *Journal of Official Statistics* 32(1): 75–92. DOI: https://doi.org/10.1515/jos-2016-0003.

Creecy, R.H., B.M. Masand, S.J. Smith, and D.L. Waltz. 1992. "Trading MIPS and memory for knowledge engineering". *Communications of the ACM* 35(8): 48–64. DOI: https://doi.org/10.1145/135226.135228.

Department for Education. 2011. *LSYPE User Guide to the Datasets: Wave 1 to Wave 7*. Available at: http://doc.ukdataservice.ac.uk/doc/5545/mrdoc/pdf/5545lsype_user_guide_wave_1_to_wave_7.pdf (accessed October 2019).

Elias, P., M. Birch, and R. Ellison. 2014. *CASCOT International version 5. User Guide*. Institute for Employment Research, University of Warwick, Coventry. Available at: https://warwick.ac.uk/fac/soc/ier/software/cascot/internat/cascot_international_user_-guide.pptx (accessed October 2019).

Gweon H., M. Schonlau, L. Kaczmirek, M. Blohm, and S. Steiner. 2017. "Three Methods for Occupation Coding Based on Statistical Learning." *Journal of Official Statistics* 33(1): 101–122. DOI: http://dx.doi.org/10.1515/JOS-2017-0006.

Hacking, W., J. Michiels, and S. Jansen, S. 2006. "Computer Assisted Coding by Interviewers." In Proceedings of the 10th International Blaise Users Conference, IBUC 2006, 9–12 May, Arnhem, The Netherlands. Available at: http://blaiseusers.org/2006/Papers/291.pdf (accessed October 2019).

Helppie-McFall, B. and A. Sonnega. 2018. *Feasibility and Reliability of Automated Coding of Occupation in the Health and Retirement Study*. Ann Arbor MI: University of Michigan Retirement Research Center (WP 2018-392). Available at: https://mrdrc.isr.umich.edu/publications/papers/pdf/wp392.pdf (accessed October 2020).

Hoffman, E. 1995. *What Kind of Work Do You Do? Data collection and processing strategies when measuring "occupation" for statistical surveys and administrative records.* ILO. (WP 1995: 95-1). Available at: https://www.ilo.org/wcmsp5/groups/public/–-dgreports/–-stat/documents/publication/wcms_087880.pdf (accessed October 2019).

Hox, J.J., E.D. De Leeuw, and E.A. Zijlmans. 2015. "Measurement Equivalence in Mixed Mode Surveys." *Frontiers in Psychology* 6(87): 1–11. DOI: https://doi.org/10.3389/fpsyg.2015.00087.

Klausch, T., B. Schouten, and J.J. Hox .2017. "Evaluating Bias of Sequential Mixed-Mode Designs against Benchmark Surveys," *Sociological Methods and Research* 46(3): 456–489. DOI: https://doi.org/10.1177/0049124115585362.

Lyberg, L., and P. Dean. 1992. *Automated Coding of Survey Responses: An International Review*. R&D Reports (1992–2). Statistics Sweden, Stockholm, Sweden. Available at:

https://www.scb.se/contentassets/7c4edb581f8745e3a081e1ba9b332eb4/rnd-report-1992-02-green.pdf (accessed October 2019).

Massing, N., M. Wasmer, C. Wolf, and C. Zuell. 2019. "How Standardized is Occupational Coding? A Comparison of Results from Different Coding Agencies in Germany." *Journal of Official Statistics* 35(1): 167–187. DOI: http://dx.doi.org/10.2478/JOS-2019-0008.

Office for National Statistics. 2010a. *Standard Occupational Classification 2010 Volume 1 Structure and descriptions of unit groups.* Available at: https://www.ons.gov.uk/-methodology/classificationsandstandards/standardoccupationalclassificationsoc/-soc2010/soc2010volume1structureanddescriptionsofunitgroups (accessed October 2019).

Office for National Statistics. 2010b. *Standard Occupational Classification 2010 Volume 2: the structure and coding index.* Available at: https://www.https://www.ons.gov.uk/-methodology/classificationsandstandards/standardoccupationalclassificationsoc/-soc2010/soc2010volume2thestructureandcodingindex (accessed October 2020).

Ossiander, E.M., and S. Milham. 2006. "A computer system for coding occupation." *American Journal of Industrial Medicine* 49: 854–857. DOI: https://doi.org/10.1002/a-jim.20355.

Schierholz, M., M. Gensicke, N. Tschersich, and F. Kreuter. 2018. "Occupation Coding During the Interview." *Journal of the Royal Statistical Society* A 181: 379–407. DOI: http://dx.doi.org/10.1111/rssa.12297.

Schierholz, M., and M. Schonlau. 2020. "Machine Learning for Occupation Coding – a Comparison Study." *Journal of Survey Statistics and Methodology, smaa023*. DOI: https://doi.org/10.1093/jssam/smaa023.

Tijdens, K. 2014. *Reviewing the measurement and comparison of occupations across Europe* (WP 149, AIAS). Available at: https://pure.uva.nl/ws/files/2172301/154005_WP149_Tijdens_1.pdf (accessed October 2019).

Tijdens, K. 2015a. The design of a tool for the measurement of occupations in web surveys using a global index of occupations. Leuven. (WP InGRID project M21.2). Available at: https://inclusivegrowth.be/downloads/output/m21-4-coding-tool-eind.pdf (accessed October 2020).

Tijdens, K. 2015b. "Self-identification of occupation in web surveys: requirements for search trees and look-up tables" *Survey Methods: Insights from the Field*. Available at: https://surveyinsights.org/wp-content/uploads/2015/06/Self-identification-of-occupation-in-web-surveys-requirements-for-search-trees-and-look-up-tables-Survey-Methods-Insights-from-the-Field-SMIF.pdf (accessed October 2020).

Tijdens, K. 2016. "*Measuring occupations: respondent's self- identification from a large database*." In Proceedings of European Conference on Quality of Official Statistics, Special session: Synergies for Europe's Research Infrastructures in the Social Sciences and Official Statistics (SERISS), 2 June 2016. Available at: https://seriss.eu/wp-content/uploads/2016/06/Measuring-Occupations-Respondent%e2%80%99s-self-identification-from-a-large-database.pdf (accessed October 2020).

Tijdens, K., and S. Visintin. 2017. *EU-harmonised and comparative measurement of occupations and skills*. Leuven. (InGRID project Deliverable 21.1). Available at: https://inclusivegrowth.be/downloads/output/d21-1-eind.pdf (accessed October 2019).

Vannieuwenhuyze, J.T.A., and G. Loosveldt. 2013. "Evaluating Relative Mode Effects in Mixed-Mode Surveys: Three Methods to Disentangle Selection and Measurement Effects," *Sociological Methods and Research* 42(1): 82–104. DOI: https://doi.org/10.1177/0049124112464868.

Vannieuwenhuyze, J.T.A., G. Loosveldt, and G. Molenberghs. 2014. "Evaluating Mode Effects in Mixed-Mode Survey Data Using Covariate Adjustment Models," *Journal of Official Statistics* 30 (1): 1–21. DOI: https://doi.org/10.2478/jos-2014-0001.