ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/136139/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Felstead, Alan 2021. Are online job quality quizzes of any value? Selecting questions, maximising quiz completions and estimating biases. Employee Relations 43 (3) , pp. 724-741. 10.1108/ER-07-2020-0349

Publishers page: http://dx.doi.org/10.1108/ER-07-2020-0349

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

ARE ONLINE JOB QUALITY QUIZZES OF ANY VALUE? SELECTING QUESTIONS, MAXIMISING QUIZ COMPLETIONS AND ESTIMATING BIASES

Alan Felstead
School of Social Sciences, Cardiff University

Purpose: This paper compares two ways of collecting job quality data in Britain using a common set of questions. One way is through a short quiz taken by a self-selected sample and completed by clicking on a web link www.howgoodismyjob.com. The other way is via an invitation to take part in a long-running survey of working life – the Skills and Employment Survey. The survey takes much longer to complete, is carried out face-to-face and is based on random probability principles.

Design/Methodology/Approach: To be content-comparable, the quiz uses tried and tested questions contained in recent waves of the Skills and Employment Survey. Each survey comprises a nationally representative sample of workers in Britain aged 20-65. However, the quiz is based on uncontrolled convenience sampling prompted, in large part, by a Facebook advertising campaign, whereas survey participants are randomly selected. In this paper, we compare the profile of respondents and their responses to these two different modes of data collection and therefore shine a light on any biases in the samples and differences in the results respondents report.

Findings: The paper shows that while the number taking in the quiz is impressive, participation in the quiz – unlike the survey – is heavily skewed. Weighting can be used to correct some of these sample selection biases. But, even then, the picture painted by the quiz and survey data varies with the quiz under-reporting the intrinsic quality of jobs, while over-reporting on the extrinsic rewards. This suggests that how job quality data are collected can have a strong influence on the results produced.

Research Implications/Limitations: The findings suggest that a number of biases are in operation, both in terms of those who take part and the answers they give. This makes comparison between data collected using radically different methods, at best, inadvisable and, at worst, misleading. Nevertheless, quizzes are a good way of engaging large numbers of people in public debates, gathering additional data, extending the reach of academic work and prompting action to improve working life. However, the limitation of this study is that it does not offer a true experiment of different ways of collecting the same data. The quiz and survey were, for example, not carried out at the same time, but were some 14 months apart.

Originality/Value: This paper focuses on how job quality data are collected and the consequences this has for the validity of the data gathered. This is a unique contribution to international debates about the measurement and monitoring of trends in job quality.

Practical Implications: Around 50,000 people took part in the quizzes reported in the paper and almost 1,300 investigated joining a trade union as a result. This far exceeds the 3,306 people who took part in the Skills and Employment Survey 2017.

Keywords: job quality; surveys; quizzes; mode effects; sampling bias; non-response bias; measurement bias; self-selection; random probability sampling; Facebook.

Paper Type: Research Paper/Technical Report
Acknowledgements

Material from the Labour Force Surveys is Crown Copyright and has been made available by the Office for National Statistics (ONS) through the UK Data Service and has been used by permission. Neither the ONS nor the UK Data Service bear any responsibility for the analysis or interpretation of the data reported here. The Skills and Employment Survey 2017 was funded jointly by the Economic and Social Research Council (ESRC), Cardiff University and the Department for Education with funding from the Welsh Government to boost the sample in Wales (ES/P005292/1). The 2012 version of the survey was funded by the ESRC/UKCES Strategic Partnership (RES-241-25-0001), with additional support from the Wales Institute of Social and Economic Research and Data. The development of the job quality quiz was funded by Cardiff University's ESRC Impact Acceleration Account (ES/M500422/1). I would also like to acknowledge the contributions of my Skills and Employment Survey colleagues, Duncan Gallie, Francis Green and Golo Henseke.

Total word count: 8,305 words (all inclusive).

Post-print manuscript, forthcoming in *Employee Relations,* 5 November 2020.

ARE ONLINE JOB QUALITY QUIZZES OF ANY VALUE? SELECTING QUESTIONS, MAXIMISING QUIZ COMPLETIONS AND ESTIMATING BIASES

Introduction

There are quicker and cheaper ways of collecting data on job quality than the random probability sample survey – online quizzes are among the quickest and cheapest of all. But do they offer a reliable way of judging whether the UK Government ambition that 'all work in the UK economy should be fair and decent with realistic scope for development and fulfilment' is being met (HM Government, 2018: 5; Taylor, 2017)?

To answer this question, this paper focuses on *how* job quality data are collected and the consequences this has for the validity of the data gathered. It compares two ways of collecting data – a random probability sample survey and an online job quality quiz. To make valid comparisons, the same questions were used for the survey and the quiz, and the data were collected within 14 months of each other and finished well before the Covid-19 pandemic began. The paper shows that while the number taking in the quiz is impressive, participation in the quiz – unlike the survey – is heavily skewed. Weighting can be used to correct some of these sample selection biases. But, even then, the picture painted by the quiz and survey data varies with the quiz under-reporting the intrinsic quality of jobs, while over-reporting on the extrinsic rewards. This suggests that how job quality data are collected can have a strong influence on the results produced. The aim of the paper is to examine this mode effect and therefore provide some cautionary evidence to those who are attracted to cheaper ways of collecting job quality data.

Research Motivation and Data Sources

The random probability sample survey is regarded by social scientists as the gold standard means of assessing societal issues. By ensuring that participants have an equal chance of being selected to take part, it is designed to be representative of the population studied. When focused on those in employment, it provides a sound statistical basis on which to estimate the quality of jobs across the population and between different socio-economic groups such as sex, age and occupation. It also has other advantages. When carried out face-to-face, for example, interviews tend to be longer than modes of data collection which do not rely on physical interaction. Furthermore, when the same questions are asked as part of a repeated series, changes in the various dimensions of job quality can be tracked and hypotheses tested using other information collected from the same respondents. The Skills and Employment Survey (SES) series is an example of this approach (Felstead *et al*., 2015). However, these surveys are relatively: expensive to carry out; infrequent; slow to produce results; subject to declining response rates; and based on small sample sizes.

Yet 'there is a strong policy need for better measures of job quality' in order to improve workers' well-being, increase productivity and competitiveness, and boost societal welfare (OECD, 2017: 14). Indicators developed for SES and research carried out using the data collected are a direct response to this need (e.g., Green, 2006; Sutherland, 2012 and 2017; Williams *et al*., 2020b; Felstead *et al*., 2020). However, the list of questions contained in existing surveys – such as SES and the CIPD UK Working Lives Survey (WLS) – is quite lengthy with SES taking 60 minutes to complete and the WLS taking 15 minutes. Furthermore, in the case of the CIPD survey those invited to take part are pre-screened and agree to complete

the survey in exchange for financial payment and entry into prize draws (YouGov, 2018). The effects that the sampling process and mode of data collection has on the results presented are not the subject of detailed discussion (Williams *et al*., 2020a: 4; Wheatley and Gifford, 2019: 3). In this paper, however, we compare the profile of respondents and their responses to two radically different modes of data collection in order to shine a light on any biases in the samples and differences in the results respondents report.

To be content-comparable, the quiz was not developed from scratch, but instead uses on tried and tested questions contained in recent waves of the Skills and Employment Survey (SES). Each of these surveys comprises a nationally representative sample of workers in Britain aged 20-65 years old at the time of data collection (although the four surveys before 2006 sampled those aged 20-60). For the 2017 survey a total of 3,306 employed individuals were interviewed in their own homes for about an hour. The samples for each of the surveys were drawn using random probability principles subject to stratification based on a number of socio-economic indicators. Only one eligible respondent per address was randomly selected for interview, and in 2017 50% of those selected completed the survey.

In this paper, SES data are used in three ways. First, data collected in 2001, 2006, 2012 and 2017 are used to demonstrate that each question selected for the quiz has a strong association with job-related well-being. This analysis is based on data taken from around 19,000 jobs. Secondly, in order to encourage participation (and answer the question 'how good is my job?') quiz takers' jobs are benchmarked against data collected in the 2012 and 2017 surveys. This amounts to data on around 6,500 jobs. Thirdly, the responses of quiz takers and survey participants are compared. In order to be as contemporaneous as possible, this analysis is

restricted to data collected in 2017 only. For this part of the paper, the analysis is based on data on 3,306 jobs and was collected between May 2017 and January 2018.

The other source of data for this paper is the data generated by the www.howgoodismyjob.com quiz itself. Data collection is still on-going, but most of it took place between July 2018-March 2019. A total of 40,125 individuals took part over this nine-month period and a further 9,295 took part in a variant of the quiz developed in association with two trade unions – the GMB and the National Union of Journalists (NUJ). This variation of the quiz ran for a month in early 2020 (www.tellmehowgoodismyjob.com). To make the data periods of the quiz and the 2017 survey as contemporaneous as possible, we focus on the quiz data collected in the period July 2018-March 2019. These two data sources provide the basis on which we make comparisons between those taking part in the quiz and the survey, and highlights the differences they gave to the questions asked.

Selecting Questions

In selecting the questions used in the job quality quiz we followed three principles: the enhancement of workers' well-being was put centre stage; all questions had to focus on features of the job; and a set of multi-faceted measures had to be used. The only restriction was that the questions selected for use in the quiz had to have been in the Skills and Employment Survey 2017 and/or the 2012 version of the same. This ensured that benchmarking feedback could be given to quiz takers and that the empirical connection between the questions selected and job-related well-being could be tested.

Ten job quality dimensions are covered by the quiz, hence making it multi-faceted. All focus on the features of the job. This means that the data are less contaminated by the personal circumstances and expectations of the individuals who do the job (see Table 1). For this reason, the quiz does not ask about respondents' anxieties over issues such as variations in pay and/or hours of work since these questions tap into subjective feelings about aspects of work (cf. Felstead *et al*., 2020).

'Put Table 1 about here'

Three of the ten dimensions can be classified as job demands since they focus on some of the pressures the work environment imposes on workers. The first focuses on the intensity of the work process as captured by the frequency with which workers report having to work at 'very high speed' and to 'tight deadlines'. The second measures the extent to which workers are required to 'keep learning new things' and are expected to help colleagues to do likewise. Job security is the third type of pressure imposed on workers. This is captured by asking quiz takers to rate their chances of job loss in the next 12 months. The remaining seven dimensions can be classified as job resources since they, potentially at least, are: 'functional in achieving work goals; reduce job demands and the associated physiological and psychological costs; [and/or] stimulate personal growth, learning, and development' (Bakker and Demerouti, 2007: 312). Quiz takers are therefore asked about: the discretion levels they are able to exercise over what tasks are to be done and how; the extent of influence they have over proposed changes to the way the job is done; the degree of control they have over starting and finishing times; the ability they have to take time off at short notice to deal with personal matters; the level of social support given by line management; their promotion prospects; and the level of pay they receive.

To confirm the link between the selected questions and job-related well-being, we pool the last four waves of the Skills and Employment Survey and examine pairwise correlations between the dimensions of job quality and well-being. The latter is measured using the Warr (1990) scales which are designed to gauge the extent to which jobs prompt arousal and pleasure. High scores are, therefore, interpreted as indicating that job-holders take great pleasure from their work and are stimulated by it, while conversely a low score is interpreted as indicating that the job is not pleasurable and fails to energise the job-holder. This analysis is based on data provided by 18,720 respondents. This confirms the existence of strong and statistically significant correlations, thereby suggesting that these features of work have the capability of enhancing or diminishing worker well-being (Felstead *et al*., 2019).

To benchmark the quality of quiz takers' jobs against similar roles, quiz takers are asked: 'What is your job title?' They are asked to respond by using words which describe their job such as 'account manager, office cleaner, web technician and delivery driver'. On typing their response, a pull-down list of similar sounding job titles appears with quiz takers selecting the most appropriate. The pull-down menu comprises 20,545 job titles provided by ONS and suitably edited and modified for quiz use (ONS, 2015). Despite the size of the database, users are warned that they may not find an exact match to their particular job title. In these circumstances, they are told to choose the closest match. This information is used to allocate quiz takers to 1-digit Standard Occupational Classification (SOC) groups.

Quiz takers were asked a series of questions on job quality. Each question relates to one of the ten job dimensions outlined above. From these data, a score ranging from 0-100 is calculated. For example, the work intensity score is derived from two questions – the frequency of working a very high speed and the frequency of working to tight deadlines. Quiz takers are given seven

response options which are scored from 0 to 6 in ascending order of frequency. These two scores are added to produce a range of 0-12 and this result is multiplied by 100/12 to produce a 0-100 range. The same principle is applied to all ten dimensions (cf. Table 1).

At the end of the quiz, a spider graph displays the results to quiz takers. These radiate out towards the perimeter for each of the 10 dimensions with the perimeter equating to a score of 100. A better job is one in which the dimension score is much higher than the occupational/overall average. Lines appear for comparator jobs as well as for all jobs. Symbols are chosen to represent the ten dimensions of job quality and are dotted around the perimeter of the plot. By moving the cursor over the symbols, the quiz taker's job is benchmarked for each of the ten dimensions. The comparator data come from the Skills and Employment Surveys in 2012 and 2017.

The movement of the cursor also triggers on-screen statements announcing the job as better, worse or about the same according to the dimension of job quality chosen. These statements appear to the left of the plot and are based on how near or far away quiz takers' responses are from the survey results. So, statements about a job being better or worse in a particular dimension only appears if the dimension score exceeds the overall or occupation averages plus or minus the margin of error (set at 10%). For comparison against similar occupations, the margins of error are typically around two points in either direction, whereas for a comparison against the British average the margin of error for most dimensions is less than one.

The quiz is designed to take, on average, five minutes to complete (15 seconds per question). Of these questions, 16 are on job quality (although two are filtered on earlier questions) and six collect information on the socio-demographic characteristics of quiz takers. The latter are

of particular use since they allow us to examine and compare the profile of quiz takers with 20-65 year olds employed and living in the Britain.[1] On this basis, we can attempt to correct for observable sampling biases.

Maximising Quiz Completions

The use of social media platforms presents unique opportunities for social scientists such as those wishing to promote a job quality quiz (Schneider and Harknett, 2019). Facebook, in particular, has two major advantages. First, its user base offers a large pool of potential research participants. Estimates suggest that in 2019 three-quarters (77%) of the UK adult population reported that they had an active Facebook profile. This equates to an estimated 37.4 million live accounts (Umpf, 2019). Secondly, the explicitly social nature of Facebook encourages users to share content which can aid 'viral' participant recruitment through likes.

To promote the quiz and maximise quiz completions, we harnessed these strengths. A variety of social media platforms – such as Facebook, Instagram and LinkedIn – were investigated but Facebook was judged to be the most effective at generating traffic to the quiz landing page and converting visitors into quiz completers. The target audience for the quiz was those aged 20-65 living in the UK and in work, making it general and wide.[2] Four separate advertising campaigns were run between July 2018 and March 2019, and direct traffic was also encouraged through two launch events held in 2018. A total of 40,125 individuals took part with a further 9,295 taking part in a variant of the quiz developed in association with the GMB and NUJ. The

[1] It is also worth noting that the quiz collects data on where people live (according to 124 postcode areas), their job titles, their 4-digit occupational group and their own qualitative assessment of their job (this includes large quantities of textual commentary).
[2] The survey focuses on Britain, while the quiz focuses on the UK. However, when making quiz-survey comparisons the analysis is restricted to Britain.

latter ran for a month in January-February 2020 (www.tellmehowgoodismyjob.com). This variant of the quiz was promoted by both trade unions through their websites, social media posts on twitter and Facebook, and newsletters to members. The survey data, on the other hand, was collected between May 2017 and January 2018. To make the data periods as contemporaneous as possible, we focus in this paper on the quiz data collected in 2018/2019.

A variety of text and images were trialled during each of the marketing campaigns. Three of these are shown in Figure 1. The success of a Facebook advert is often judged in terms of its 'cost per click'; that is, the average cost spent per Facebook user who clicks onto the advert in response to the call for action. The conversion rate refers to those who land on the site and subsequently carry out a further action, in this case complete the quiz. The most successful advert – measured in terms of the cost per click – was an advert comprising text quick call to action to find out 'how good is your job' accompanied by stylised images of a variety of jobs and a university logo (see Figure 1a). The cost per click for the 2018/2019 quiz was 9.5p which varied little by age and sex, and suggests that the advert was widely seen across all ages and sexes. A total of 531,726 Facebook users saw the advert and it appeared in their News Feed an average of 2.4 times. The click through rate was 10.7% and the completion rate was 7.5%, meaning that each quiz completion cost 13.5p to secure.

'Put Figure 1 about here'

However, the major disadvantage of Facebook is that despite its size, its user pool does not offer a representative sampling frame for social science research. For example, women are more likely than men to have an active Facebook profile – 79.6% of women are Facebook users compared to 71.9% of men. There are also age divides with nearly all (94.8%) of 18-25 year

olds actively using Facebook compared to 63.5% of those aged 55-64. Although there is evidence that the popularity of Facebook among the young is fading while its popularity among the older generation is increasing, a sharp age divide nonetheless remains (e.g., *The Guardian*, 12 February 2018). Frequency of use – and therefore exposure to advertising – differs in a similar fashion with, for example, a fifth of women (21.2%) report using Facebook more than five times a day compared to a sixth of men (17.5%) (Umpf, 2019). However, other forms of social media are skewed, but in different ways. Twitter, for example, is disproportionately used by men and is much more the preserve of the young (Sloan, 2017).

Estimating Biases

Sampling and Non-Response Biases

To estimate biases in both the quiz and survey samples, we use the relevant Labour Force Survey (LFS) as the comparator. Despite keen policy interest in data on the labour market, there is no national sampling frame of jobs. So, like other sample surveys, the latest version of the small user Postcode Address File was used to provide a list of households from which to sample for SES. Postcode sectors were placed in a stratified list according to the size of the sub-region, the number employed in higher level occupations and the number unemployed. From this list, postcode sectors were selected using a random start point and moving at a fixed interval to select the required number of postcode sectors. From these sectors, addresses were randomly selected.

When interviewers made contact with sampled addresses they had to select one dwelling, then one household and finally one eligible individual within the selected household. In the vast

majority of cases, no dwelling or household selection was necessary, but where multiple dwellings or households or eligible individuals were found, interviewers used a 'kish grid' to select at random one eligible respondent. The 'Kish grid' is a list of random numbers generated according to the number of dwellings, households and eligible respondents. This ensures that the selections are made at random.

Despite these efforts, the unweighted sample differs a little from the LFS estimates of the socio-economic composition of adults aged 20-65 and working in Britain. We chose the LFS since it is based on a large sample of households with almost 40,000 households contacted every quarter and around 42,000 workers interviewed. On this basis, the LFS sample is likely to be closely representative of the employed workforce. A comparison of the unweighted SES2017 with the LFS for the second quarter of 2017 suggests that the former under-represents men, the young, higher occupational groups and those working in the private sector (see Table 2).

'Put Table 2 about here'

Corrections are, therefore, required. By design, unequal selection probabilities can occur at three points in the selection process. When selecting: one dwelling per address; one household per dwelling; and one eligible adult per household. In many datasets, these are referred to as 'Kish weights'. Where there are no boosts to the sample sizes, Kish weights can be used to produce the sample distributions. However, in the case of the 2006, 2012 and 2017 Skills and Employment Surveys, the samples for certain areas were boosted and so a weight was derived to take this into account. So, given that Wales was over-sampled in all three surveys Welsh respondents were given a weight of less than one to reflect this intentional over-sampling (note the over-representation of Welsh respondents in the unweighted sample in Table 1).

Corrections are also needed to adjust for non-response as seen in the over- and under-representation of certain groups in the unweighted sample. This is achieved by making the weights of each of the groups in Table 1 inversely proportional to the LFS estimates (shown in column 2, Table 2). These group weights along with the Kish weight (discussed above) are then multiplied to produce a survey weight. When applied to the SES2017 data, the profile of the sample moves closer the LFS estimates, thereby making the weighted results more representative.

The quiz is also biased by the Facebook marketing campaign and the events which publicised the headline findings of the 2017 survey. Both of the quiz sampling frames are biased in two ways: first, in terms of who saw the publicity for the quiz and secondly, who then decided to take part. The resulting sample is therefore skewed and is therefore unlikely to be representative of the jobs undertaken by adults.

The unweighted sample demonstrates the extent of the problem. Like the survey, women are over-represented in the quiz, but on a much larger scale – two-thirds of the quiz sample are female (66.1%).[3] The young are also over-represented in the quiz sample as are those in higher skilled jobs and those working in the public sector. Again, the over-representation is marked.

'Put Table 3 about here'

[3] Similar gender biases are found in other online surveys using uncontrolled sampling. For example, Eurofound's Living, Working and Covid-19 quiz was promoted through social media. Around 70% of its respondents were female (Sandor and Ahrendt, 2020).

We attempt to correct for these sampling biases by creating a quiz weight. This is achieved by following the same principles outlined above; that is, by making the weights of each of the groups in Table 3 inversely proportional to the LFS estimates (shown in column 2, Table 3) and then multiplying them together to produce a quiz weight. When applied, the profile of the quiz moves closer to that of the LFS. However, large gaps remain. Most glaringly, after weighting the quiz men become over-represented in the sample by five percentage points, while they are under-represented by twenty points in the unweighted sample. However, the weighted quiz sample is representative of six out of the eleven regions/countries.[4] That said, the survey performs even better in that it is representative of all regions/countries and all socio-economic indicators, apart from economic sector.[5] This suggests that while the survey data is representative of the adult working population in Britain, the quiz data – even after weighting – is not a representative sample.

Measurement Biases

To make matters worse, there are other biases which may influence how survey respondents and quiz takers respond to the questions posed. Even though the *same questions* are asked of different samples of workers (as in the quiz and survey reported here), the mode of data collection may have an effect on the answers given.

There are several reasons for this. The first reason is that the quiz itself is intended to prompt curiosity with the URL www.howgoodismyjob.com. It makes a direct appeal to people's

[4] Statistical tests were carried out which compared the weighted survey and quiz samples with their respective LFS comparators by dividing the difference between the estimated means of each socio-economic characteristic by the square root of the summed squared standard errors.
[5] However, in line with the weights used for other surveys in the SES series (and supplied with the publicly available datasets) the survey weight for 2017 does not correct for sector differences.

inquisitive nature, especially since there are few other means of comparing one's job with those of others. That said, there online tools which provide pay comparisons. These tools allow users to compare their pay against others working in a similar position and locality, and therefore better assess what their labour is worth (e.g., www.glassdoor.co.uk and www.totaljobs.com). However, these sites tend to be targeted at those seeking to move jobs and putting these individuals in contact with employers seeking new recruits. The www.howgoodismjob.com quiz, on the other hand, directly appeals to those who are more curious and are therefore more likely to complete the quiz. This may help to explain the skewed nature of the sample compared to the survey where participation is on an invitation-only basis. It may also influence the nature of the data collected since curiosity-motivated quiz takers tend to have a more positive outlook compared to those invited to take part in a survey. They may also take part in order to 'show off' (share) their results on social media and therefore exaggerate their extrinsic job rewards (Kashdan *et al*., 2020).

However, social desirability bias may operate in the opposite direction. This is the second reason why results may differ between the two different modes of data collection. It is known that respondents tend to react differently in surveys using face-to-face interviewing compared to quizzes completed online. This is because respondents are more reluctant to report socially undesirable traits and negative features of their lives to an interviewer. In other words, survey respondents tend to choose positive responses over negative ones on the basis that these responses are more socially desirable rather than responses which reflect their lived experience (DeMaio, 1984). This is especially the case when interviews are carried out in the presence of other household members as can often happen when surveys, such as SES, are based on household sampling. Online quizzes, on the other hand, are less susceptible to social

desirability bias and may therefore generate results which are closer to respondents' lived experience.

Similarly, the theory of satisficing suggests that quiz takers may be more motivated and willing to engage in the cognitive work required to give honest and careful answers to a five-minute quiz. Whereas survey respondents may be less thorough in giving their answers in interviews lasting an average of 60 minutes. In other words, survey respondents may satisfice the interviewer by giving responses which sound reasonable rather than thinking carefully about the responses they give (Krosnick, 1991).

To examine how responses to the two modes of data collection differ, we contrast the answers given to the same questions asked of the two groups (see Table 4). In descriptive terms (and after weighting), quiz takers offer responses which suggest that their jobs are poorer on six out of the ten job quality dimensions, better on three dimensions and about the same as their survey counterparts on one. For example, over one in six quiz takers (17.2%) report that they have an evens or higher chance of losing their job in the next 12 months compared to around one in ten (9.2%) of those who took part in SES2017. Similarly, quiz takers report lower levels of line management support than survey participants as indicated by a significantly lower managerial support score. However, quiz takers rate their promotion prospects as higher and report higher pay than survey participants. These comparisons are made after the samples have been weighted with the result that the differences reported in Table 4 narrow and widen as expected as weights are turned on and off. Given the over-representation of higher occupational groups in the unweighted quiz sample, quiz-survey differences on dimensions – such as insecurity, managerial support and the requirement to learn – widen when weighting is applied. On the

other hand, other differences – such as pay, work intensity, job discretion and voice – narrow when the data are weighted.

'Put Table 4 about here'

As an additional test, we carry out a series of weighted regressions using the ten job quality dimension scores as dependent variables. These use all the control variables at our disposal, but they are limited to questions asked in the quiz. Nevertheless, we control for gender, age, sector, occupation and region. The results largely confirm the bivariate analysis with quiz takers reporting that their jobs are significantly poorer than survey participants on six out of ten measures of job quality, about the same in terms of discretion and voice, and significantly better in terms of pay and promotion. Furthermore, these differences persist within occupational groups. For example, teachers and health-related professionals who took part in the quiz reported higher pay and better promotion prospects than their counterparts who responded to the survey.[6] Overall, then, these patterns suggest that the reporting biases outlined above operate differently according to the nature of the data reported. In particular, features of intrinsic job quality tend to be under-reported by quiz takers, while extrinsic features such as pay and promotion tend to be over-reported.

'Put Table 5 about here'

[6] We examine 'teachers' defined as those in SOC 4-digit categories (2314 to 2319) and 'health-related professions' such as nurses and midwives (defined by the 4-digit categories of 2211 to 2232). We focus on these are two occupational categories since we have a reasonable number of observations in both the quiz and the survey (over 2,000 cases in the quiz and 125 or more in the survey). They also form part of the 'professional' group which is over-represented in the quiz, therefore skewing the results if different patterns are found.

However, despite identifying these differences, we have no way of adjusting these data to take into account these mode effects. In other words, unlike calculating sampling and non-response weights (as outlined earlier) in order to adjust for the effects of unequal sample selection, an equivalent mode effect adjustment process is not available. Nevertheless, the scale of the disparities suggests that social desirability and curiosity biases operate differently when job quality data are collected by an online quiz compared to a face-to-face survey.

Conclusion

Before the Covid-19 pandemic, job quality was making headlines with policy makers at all levels – internationally, nationally and regionally – keen to promote good jobs by using whatever means at their disposal. There has, for example, been a spate of local initiatives designed to promote good work in particular localities. In the UK these have often been driven by Labour controlled authorities in large urban conurbations such as London, Manchester and Liverpool (Hurrell *et al*., 2017). The devolved administrations have also developed pledges and codes designed to drive good employment practice (Welsh Government, 2017; Scottish Government, 2019). Trade unions, too, have launched charters as a campaigning device to highlight employment relations issues in particular sectors. In 2018 Unite issued a Construction Charter as well as a Fair Hospitality Charter, while Unison launched its Ethical Care Charter in 2012 with the aim of promoting both good standards of domiciliary care and good employment practices (Moore, 2017).

This paper assesses a new and novel way of collecting data job quality – the online quiz. Unlike the traditional survey, the quiz harnesses the power of social media to connect with large numbers of people through everyday devices such as the mobile phone, the tablet and the

computer. The paper has two key findings. First, quiz participation is skewed towards women and younger workers, and those working in higher skilled jobs and the public sector. Furthermore, sampling and non-response rate adjustments fail to produce a representative quiz sample since the skewed nature of the original sample is simply too great. Survey participation, on the other hand, is – even before weighting – more representative of the adult working population in Britain. This provides a robust basis on which to track movements in job quality and assess its drivers. Secondly, the substantive results produced by these two particular methods of data collection differ significantly. This suggests that factors such social desirability bias are at work, making comparison between data collected using radically different methods, at best, inadvisable and, at worst, misleading (Bowling, 2005).

However, the limitation of this study is that it does not offer a true experiment of different ways of collecting the same data. The quiz and survey were, for example, not carried out at the same time, but were some 14 months apart. It is also possible that the quiz was completed on multiple occasions by the same individual, whereas participation in the survey was a one-time only event. Even so, the samples and the substantive results of the quiz and the survey are so different that it is difficult to conclude that they are commensurate. For these reasons, Eurofound, for example, has refrained from making direct comparisons between its recently completed Covid-19 online quiz and data collected from the European Working Conditions Survey, even when both samples were asked exactly the same question. Data from the latter are only provided 'as a source of information' and not as a point of comparison (Eurofound, 2020: 7).

What, then, are the benefits of job quality quizzes? First, they promote public engagement with an issue of policy relevance. The quiz reported in this paper has engaged almost 50,000

individuals in a conversation about job quality. Quiz takers have also, unwittingly, engaged with one of the largest and most frequently carried out surveys of job quality in Britain, hence generating additional publicity for its results. Secondly, quiz takers have been able to benchmark their jobs against others in a way previously not possible. Thirdly, the quiz has generated a vast amount of data collected from many more people than typically take part in surveys. While participation is skewed, the resulting data allows patterns to be observed at much finer levels of granularity, such as by detailed occupation and by postcode area, than is possible using surveys (Williams *et al*., 2020b). Fourthly, variants of the quiz have been promoted by other organisations, such as trade unions, to highlight the value of their services. For example, out of around 9,500 who took the trade union variant of the quiz (www.tellmehowgoodismyjob.com), almost 1,300 investigated joining a trade union as a result.[7] Job quality quizzes can, therefore, prompt individuals into taking action on the basis a quiz. However, none of these benefits extends to the provision of robust, time series data which – while costly and time-consuming – remains the preserve of random probability sample surveys carried out at regular intervals.

[7] Non-union members are invited to learn more about joining a relevant trade union by clicking to find out more and going to the relevant trade union's 'join-up' page. Using Google Analytics, the number of clicks made can then be counted.

References

Bakker, A. B. and Demerouti, E. (2007) 'Job Demands-Resources Model: state of the art', *Journal of Managerial Psychology*, 22(3): 309-328.

Bowling, A. (2005) 'Mode of questionnaire administration can have serious effects on data quality', *Journal of Public Health*, 27(3): 281-291.

DeMaio, T. J. (1984) 'Social desirability and survey measurement: a review', in Turner, C. F. and Martin, E. (eds) *Surveying Subjective Phenomena*, New York: Russell Sage.

Eurofound (2020) *Living, Working and COVID-19*, Luxembourg: Publications Office of the European Union.

Felstead, A., Gallie, D. and Green, F. (2015) (eds) *Unequal Britain at Work*, Oxford: Oxford University Press.

Felstead, A., Gallie, D., Green, F. and Henseke, G. (2019) 'Conceiving, designing and trailing a short form measure of job quality: a proof-of-concept study', *Industrial Relations Journal*, 50(1): 2-19.

Felstead, A., Gallie, D., Green, F. and Henseke, G. (2020) 'Unpredictable times: the extent, characteristics and correlates of insecure hours of work in Britain', *Industrial Relations Journal*, 51(1-2): 34-57.

Green, F. (2006) *Demanding Work: The Paradox of Job Quality in the Affluent Economy*, Princeton: Princeton University Press.

HM Government (2018) *Good Work Plan*, London: Department for Business, Energy and Industrial Strategy.

Hurrell, D.-L., Hughes., C and Ball, E. (2017) *Local Employment Charters: Case Studies from the UK*, Manchester: Inclusive Growth Analysis Unit.

Kashdan, T. B., Goodman, F. R., Disabato, D. J., McKnight, P. E., Kelso, K. and Naughton, C. (2020) 'Curiosity has comprehensive benefits in the workplace: developing and validating a multidimensional workplace curiosity scale in the United States and German employees', *Personality and Individual Differences*, 155(1): 109717.

Krosnick, J. A. (1991) 'Response strategies for coping with the cognitive demands of attitude measures in surveys', *Applied Cognitive Psychology*, 5(3): 213-236.

Moore, S. (2017) *An Evaluation of UNISON's Ethical Care Charter*, London: University of Greenwich.

OECD (2017) *OECD Guidelines on Measuring the Quality of the Working Environment*, Paris: Organisation for Economic Co-operation and Development.

ONS (2015) *The Coding Index of SOC2010*, Newport: Office for National Statistics.

Sandor, E. and Ahrendt, D. (2020) *Living, Working and COVID-19: Methodological Annex to Round 2*, Dublin: European Foundation for the Improvement of Living and Working Conditions.

Schneider, D. and Harknett, K. (2019) 'What's to like? Facebook as a tool for survey data collection', *Sociological Methods and Research*, online first.

Scottish Government (2019) *Scottish Business Pledge: Proposal Form,* Glasgow: Scottish Government.

Sloan, L. (2017) 'Who tweets in the United Kingdom? Profiling the Twitter population using the British Social Attitudes Survey 2015', *Social Media + Society*, 3(1): 1-11.

Sutherland, J. (2012) 'Job attribute preferences: who prefers what?', *Employee Relations*, 34(2): 193-221.

Sutherland, J. (2017) 'Employee well-being in Britain: perspectives from the margins of the labour market', *International Journal of Social Economics*, 44(12): 1-19.

Taylor, M. (2017) *Good Work: The Taylor Review of Modern Working Practices*, London: Department for Business, Energy and Industrial Strategy.

The Guardian (2018) 'Is Facebook for old people? Over-55s flock in as the young leave', by Sweney, M, 18 February.

Umpf (2019) *Social Media Usage in the UK – The Findings*, Leeds: Umpf, https://umpf.co.uk/social-media-usage-in-the-uk-the-findings/, accessed 14 June 2020.

Warr, P. (1990) 'The measurement of well-being and other aspects of mental health', *Journal of Occupational Psychology,* 63: 193-210.

Welsh Government (2017) *Code of Practice: Ethical Employment in Supply Chains*, Cardiff: Welsh Government.

Wheatley, D. and Gifford, J. (2019) *UK Working Lives: The CIPD Job Quality Index, Appendix 1*, London: Chartered Institute of Personnel and Development.

Williams, M., Zhou, Y. and Zou, M. (2020a) *CIPD Good Work Index 2020: UK Working Lives Survey*, London: Chartered Institute of Personnel and Development.

Williams, M., Zhou, Y. and Zou, M. (2020b) *Mapping Good Work: The Quality of Working Life Across the Occupational Structure*, Bristol: Policy Press.

YouGov (2018) 'Our panel', https://yougov.co.uk/about/our-panel/, accessed 10 June 2020.

Table 1:
Job Quality Dimensions: Questions, Scales and Content Validity

Job Quality Dimension	Question(s)	Scale	Content and Construct Validity[1]
1. Job insecurity	Do you think there is any chance at all of you losing your job and becoming unemployed in the next twelve months? [If yes] How would you rate the likelihood of this happening?	if no to first question=0; very unlikely to second follow-on question=1; quite unlikely to second follow-on question=2; evens to second follow-on question=3; quite likely to immediate second follow-on question=4; and very likely to second follow-on question=5. There are two questions, but one filtered as a follow-on, so the score ranges from 0-5.	NA – single item filtered; r=0.18***
2. Working time autonomy	How much do you agree or disagree with the statement 'I can decide the time I start and finish work'?	strongly disagree=0; disagree=1; agree=2; and strongly agree=3 (i.e., 0-3).	NA – single item; r=0.11***
3.Work-life balance	Would you say that for you arranging to take an hour or two off during working hours to take care of personal or family matters is….?	very difficult=0; somewhat difficult=1; not too difficult=2; and not at all difficult==3 (i.e., 0-3).	NA – single item; r=0.16***
4.Manager-ial support	How helpful is your supervisor or manager in recognising the extent of your abilities? How helpful is your supervisor or manager in enabling you to learn how to do your job better?	of no help at all=0; a little help=1; of some help=2; quite a lot of help=3; and a great deal of help=4 (i.e., 0-4).	α=0.69; r=0.34***
5. Required learning	How much do you agree or disagree that my job requires that I learn new things? How much do you agree or disagree that my job requires that I help my colleagues learn new things?	strongly disagree=0; disagree=1; agree=2; and strongly agree=3 (i.e., 0-3).	α=0.69; r=0.12***
6. Intensity	How often does your work involve working at very high speed? How often does your work involve working to tight deadlines?	never=0; almost never=1; around a quarter of the time=2; around half the time=3; around three-quarters of the time=4; almost all the time=5; and all the time=6 (i.e. 0-6) for both questions.	α=0.74; r=0.11***
7. Discretion	How much influence do you *personally* have on deciding what tasks you are to do?	none=0; not much=1; a fair amount=2; and a great	α=0.77; r=0.22***

	How much influence do you *personally* have on deciding on how you are to do the task?	deal=3 (i.e., 0-3) for both questions.	
8.Voice	Suppose there was going to be some decision made at your place of work that changed the way you do your job. Do you think that you personally would have any say in the decision about the change or not? [If yes] How much say or chance to influence the decision do you think that you personally would have?	no to the first question=0; it depends to first question=1; just a little to second follow-on question=1; quite a lot to second follow-on question=2; and a great deal to second follow-on question=3 (i.e., 0-3).	NA – single item filtered; r=0.27***
9.Promotion prospects	Assuming that you wanted promotion, how high do you think *your* chances are of being given a significant promotion with your *present* organisation in the next five years?	no chance at all=0; 25%/low chance=1; 50%/fifty-fifty=2; 75%/high chance=3; and a 100%/definite=4 (i.e., 0-4).	NA – single item; r=0.14***
10. Pay	What is your usual pay including overtime, bonuses or tips (but before tax and other deductions are taken out)?	1=£7.49 an hour or less (£14,600 per year or less for a full-time job); 2=£7.50 - £10.00an hour (£14,601 - £19,500 per year for a full-time job); 3=£10.01 - £15.00 an hour (£19,501 - £29,250 per year for a full-time job); 4=£15.01 - £21.00 an hour (£29,251 - £40,950 per year for a full-time job); 5=£21.01 - £31.00 an hour (£40,951 - £60,450 per year for a full-time job); 6=£31.01 an hour or over (£60,451 per year or over for a full-time job). The SES2017 hourly pay data use equivalent bands. The pay banding score ranges from 1-6.	NA – single item in quiz, pay per hour data in SES; r=0.03**

Notes:

1. Cronbach's alphas (α) are based on all the available data for each of the items in the seven surveys which make up the Skills and Employment Survey. This ranges from 24,605 for task discretion (that is, the 1992, 1997, 2001, 2006, 2012 and 2017 surveys) to 2,793 for managerial support (the 2017 survey only). To capture job-related enthusiasm, respondents to the 2001, 2006, 2012 and 2017 Skills and Employment Surveys were asked: 'Thinking of the past few weeks, how much of the time has your job made you feel each of the following…?' Respondents were asked about a series of adjectives, each describing a different feeling. To construct an enthusiasm scale we use responses given to the following the adjectives: 'depressed', 'gloomy', 'miserable', 'cheerful', 'enthusiastic' and 'optimistic'. The response set comprised six points ranging from 'never' to 'all of the time'. We construct a scale by averaging the responses given (after reversing the three negative items) and using a 1-6 scoring system. These questions were only asked from 2001 onwards, so the results in this column are restricted accordingly. Pairwise correlations (r) are presented. * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Source: adapted from Felstead et al., 2019: Table A1.

Figure 1:
Facebook Advertising: Text and Images Used

(a) Simplistic Text Emphasising Speed and Vector Images of Jobs and University Logo	*(b) Detailed Text On Evidence Base for Comparisons and Image Showing a Mixed Group of Real Workers*	*(c) Less Detailed Text Emphasising Research and Vector Images of a Variety of Jobs*

Table 2:
Sampling and Non-Response Biases: Skills and Employment Survey 2017

Demographics and Socio-Economic Characteristics	Unweighted	Labour Force Survey[1]	Weighted
Gender			
Male	48.6	53.0	53.1
Female	51.4	47.0	47.0
Age			
20-29	15.7	22.0	21.9
30-39	24.7	24.1	24.3
40-49	25.3	24.1	24.2
50-60	26.1	24.6	23.5
61-65	8.2	5.2	6.1
Occupation			
Managers, Directors and Senior Officials	9.0	10.8	10.7
Professionals	19.0	20.8	20.5
Associate Professionals	16.1	14.8	14.7
Administrative & Secretarial	8.7	10.4	9.8
Skilled Trades	11.2	10.5	11.2
Caring and Leisure	11.0	9.4	8.9
Sales and Customer Service	7.2	7.2	7.5
Plant & Machine Operatives	7.4	6.5	6.9
Elementary	10.4	9.6	9.8
Region			
North East	4.0	4.0	4.1
North West	10.0	11.1	10.6
Yorkshire and the Humber	8.3	8.2	8.7
East Midlands	6.5	7.2	7.2
West Midlands	7.4	8.6	8.1
East of England	9.7	9.7	11.6
London	12.2	15.2	15.2
South East	12.1	14.5	14.1
South West	8.0	8.6	8.4
Wales	13.8	4.6	4.5
Scotland	8.0	8.6	7.5
Sector			
Private sector	64.8	72.9	67.3
Public sector	30.9	22.4	28.3
Non-profit organisation	4.3	4.7	4.4

Note:

1. These data are based on the most contemporaneous Labour Force Survey (Q2, 2017) to the time when most of the SES interviews were completed.

Table 3:
Sampling and Non-Response Biases: Job Quality Quiz 2018-2019

Demographics and Socio-Economic Characteristics	Unweighted	Labour Force Survey[1]	Weighted
Gender			
Male	32.6	52.7	57.9
Female	66.1	47.3	41.0
Other	0.2	NA	0.2
Prefer not to say	1.1	NA	1.0
Age			
16-19	0.4	3.3	0.3
20-29	29.4	20.7	22.9
30-39	25.4	23.7	24.4
40-49	20.8	22.9	21.8
50-60	18.4	22.5	23.7
61-65	4.5	7.0	5.9
Prefer not to say	1.1	NA	1.0
Occupation			
Managers, Directors and Senior Officials	13.2	11.1	12.5
Professionals	34.3	22.0	18.9
Associate Professionals	17.0	14.8	14.6
Administrative & Secretarial	10.8	9.9	8.2
Skilled Trades	3.9	10.1	13.7
Caring and Leisure	9.0	9.1	6.8
Sales and Customer Service	5.1	6.9	6.5
Plant & Machine Operatives	2.4	6.3	9.1
Elementary	4.3	9.6	9.6
Region			
North East	3.6	3.7	3.8
North West	11.0	10.8	10.6
Yorkshire and the Humber	8.9	8.0	8.1
East Midlands	5.9	7.1	7.1
West Midlands	8.3	8.5	8.6
East of England	9.8	9.5	9.1
London	8.2	14.7	12.2
South East	12.8	13.9	13.5
South West	10.2	8.4	8.0
Wales	5.9	4.6	4.2
Scotland	9.0	8.3	8.0
Northern Ireland	1.6	2.6	2.6
Not known	5.0	NA	4.3
Sector			
Private sector	57.2	73.0	78.3
Public sector	35.1	22.6	17.8
Non-profit organisation	7.7	4.5	3.9

Note:

1. These data are based on three Labour Force Surveys (Q3 2018, Q4 2018 and Q1 2019) which cover the period when most of the quizzes were completed.

Table 4:
Measurement Bias: Dimensions of Job Quality, Bivariate Comparisons

Dimensions of Job Quality[1]	Skills and Employment Survey 2017 (% and scores)	Job Quality Quiz 2018-2019 (% and scores)
1. Job Insecurity		
Evens or higher chance of losing job in next 12 months	9.2	17.2
Job insecurity score[2]	0.42	0.74***
2. Working Time Autonomy		
Strong disagreement that 'I can decide the time I start and finish work'	27.5	35.7
Working time autonomy score	2.35	2.17***
3. Work-life Balance		
Very difficult to take time off to take care of personal or family matters	11.1	19.4
Work-life balance score	1.91	1.65***
4. Managerial Support		
Line manager is a great deal of help in recognising the extent of abilities	26.2	15.8
Line manager is a great deal of help in enabling learning	27.5	20.5
Managerial support score	2.71	2.18***
5. Required Learning		
Disagreement that job requires keep learning new things	15.7	20.4
Disagreement that job requires helping others to learn new things	16.2	22.4
Required learning score	3.17	3.07***
6. Intensity		
Working at very high speed a half or more of the time	66.4	70.3
Working to tight deadline a half or more of the time	75.0	76.9
Intensity score	4.54	4.64***
7. Discretion		
None or not much influence over what tasks are to be done	28.2	33.5
None or not much influence over how to do the tasks	16.9	19.1
Discretion index score	2.08	2.06
8. Voice		
A great deal of say over decisions to change the way the job is done	11.9	12.7
Voice score	1.05	1.09*
9. Promotion Prospects		
Definite/100% chance of being promoted	5.2	10.9
Promotion prospects score	2.27	2.62***
10. Pay		
Take home pay less than £7.49 an hour	14.2	4.2
Take home pay more than £31.01 an hour	5.8	9.3
Pay banding score	3.01	3.42***

Notes:

1. The data presented here and in Table 5 restrict the quiz sample to those 20-65 and living in Britain in order to be comparable to the Skills and Employment Survey 2017. Weights are applied to the survey and the quiz.
2. Scores are calculated as outlined in Table 1, with averages taken where necessary. Significance tests compare the quiz score with the survey score, * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 5:
Measurement Bias: Dimensions of Job Quality, Weighted Regressions

Independent Variables	Dimensions of Job Quality as Measured by Scores[1]									
	Job Insecurity (1)	Working Time Autonomy (2)	Work-life Balance (3)	Managerial Support (4)	Required Learning (5)	Intensity (6)	Discretion (7)	Voice (8)	Promotion Prospects (9)	Pay (10)
Quiz (survey is reference)	0.299*** (0.027)	-0.099*** (0.020)	-0.246*** (0.020)	-0.527*** (0.023)	-0.095*** (0.014)	0.098*** (0.028)	0.013 (0.015)	-0.006 (0.019)	0.053*** (0.004)	0.290*** (0.023)
Controls[2]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.016	0.124	0.086	0.033	0.131	0.024	0.116	0.081	0.134	0.386
Number of observations	41,237	41,495	41,477	41,415	41,403	41,480	41,491	41,465	41,371	40,990

Note:

1. Scores are used as summary measures of the range of responses given. For detail, see Table 1 and footnote 2 in Table 4.
2. Controls are a bivariate dummy for gender, four dummies for age, eight dummies for jobs categorised according to 1-digit Standard Occupational Classification system, two dummies for sector and eleven dummies for region.