

What social media can tell us about essential occupations

Contextualising twitter data to understand shifts in occupational valuations


Dr. Michael Tiemann ¹, Stefan Udelhofen ² and Lisa Fournier ³


Abstract: Societal debates about essential occupations in the context of the pandemic have raised questions about the valuation of occupational tasks. In a first step we compare two lists of essential occupations, one from the start, the other from the end of the pandemic, to describe differences in their valuation based on characteristics such as wages, prestige and workload. Between these lists it becomes apparent that there has been a broadening shift, with essential occupations at the end of the pandemic being different to what they were at the beginning. This is based on data from the BIBB/BAuA Employment Survey 2018. We then investigate the use of twitter data for generating insights on how the valuation of occupations were discussed and changed during the pandemic and thus helped leverage said shift in essential occupations.


Keywords: essential occupations, COVID-19, Twitter

1 Essential occupations under discussion⁴

Occupations relating to the provision of key supplies and services – otherwise known as “essential jobs” or “essential work” – are still new categories in the field of labour market and occupational research [HSZ23]. Different occupational groups are subsumed under this heading depending on the context of a specific crisis situation, although no clear definition for essential occupations exists. They gained some attention in the light of the first coronavirus lockdown in March 2020. In Germany, the Federal Office for Civil Protection and Disaster Assistance categorised them as occupations which “help ensure provision to the population of important and in some cases vital goods and services” (own translation, cited from [SSS21]). Workers in these occupations enjoyed special rights (e.g. emergency childcare), but were also subject to restrictions (such as bans on taking a leave

¹ Federal Institute for Vocational Education and Training, AB 12, Friedrich-Ebert-Allee 114-116, 53113 Bonn, tiemann@bibb.de,  <https://orcid.org/0000-0001-7136-2744>

² Federal Institute for Vocational Education and Training, AB 12, Friedrich-Ebert-Allee 114-116, 53113 Bonn, stefan.udelhofen@bibb.de,  <https://orcid.org/0000-0002-1749-7231>

³ Federal Institute for Vocational Education and Training, AB 12, Friedrich-Ebert-Allee 114-116, 53113 Bonn, lisa.fournier@bibb.de,  <https://orcid.org/0009-0008-9544-5577>

⁴ The parts of this paper dealing with the analysis of BIBB/BAuA-Employment Survey data (section 1) are in large parts based on our article [TUF23].

and the lack of opportunities for working from home). These “initial” [Ko20]⁵ lists of essential occupations were adjusted and extended as the pandemic continued. Teaching staff and financial services workers, for example, were added. Regional differences between individual federal states also emerged, such as in how agricultural occupations were considered.

Our article seeks to take this as a starting point in order to link a description of the valuation of essential occupations with the question of the distribution of requirements and qualification levels. A consideration of the categorical dynamics produced via the adjustment of lists of essential occupations appears to be of particular interest, also extending the focus of previous studies. For this reason, we use data from the 2018 BIBB/BAuA Employment Survey⁶. The Classification of Occupations (KldB 2010) is used to operationalise and differentiate between two lists: a narrower list of occupations (referred to below as the “Berlin list”), which were defined on an ad hoc basis as being essential in the wake of the coronavirus pandemic (20 occupations), and an extended list (62 occupations), which takes account of subsequent additions made by policy makers and of the findings from academic research studies.

In our research, we looked at characteristics associated with valuation. For this, we have looked at pay, job prestige, workload and demands, and skill levels of essential jobs. We also investigated whether there is a mismatch between the requirement levels of the tasks and the qualification levels of the employees.

As indicators of occupational esteem, we examined pay and prestige. Initially, our comparative investigation of remuneration and occupational prestige as indicators of the valuation and esteem of individual occupations is essentially in line with findings from previous studies. In the essential occupations on the “Berlin list”, the average hourly wage is significantly lower (€16.32) than that paid in non-essential occupations (€19.22). This discrepancy is not reproduced in the “extended list”. This is similar with regard to occupational prestige. The average ISEI value⁷ for occupations on the “Berlin list” (40.15) is significantly below the value of the other occupations (47.88). Although the average value for occupations on the “extended list” is slightly above that for all other occupations (47.05 as compared to 44.96), this difference is not significant.

⁵ Actually, there have been several lists from the beginning due to most federal states issuing their own list.

⁶ The Employment Survey 2018 is a representative survey of 20,012 employees in Germany on changes in work and occupations and on the acquisition and utilisation of professional and vocational qualifications. The data [HHR18a; HHR18b] were collected by Kantar Public of Munich via computer-assisted telephone interviews (CATI) carried out between October 2017 and April 2018. Selection was based on a random procedure (Gabler-Häder sampling process) in order to ensure representativity. The statistical population comprises members of the labour force aged 15 and above (not including trainees). Employment activity is considered to be regular work activity of at least ten hours per week for which payment is received (“core workers”). Data was adapted to the structures of the population via weighting in accordance with central characteristics on the basis of the 2017 Microcensus.

⁷ The “International Socio-Economic Index of Occupational Status” (ISEI) maps the social status of an occupation by providing a value between 16 and 90 and thus ranks occupations according to status. Further information is available at: <https://metadaten.bibb.de/de/classification/detail/11>.

There are other measures to take into account when looking at valuations. In the context of essential occupations the subjective valuation of ones own work should be taken into account as well. Here it is worthwhile to note that employees in occupations on both lists are similarly likely to rate their work as important. Differences between occupations on the two lists are shown with regard to consideration of the employees' situation. Workers in occupations on the "Berlin list" tend to evaluate their overall work conditions as being worse and also state that workloads are higher than average. They are also more likely to be less well informed about tasks they have to fulfil and report that they enjoy a lesser degree of occupational autonomy (for more details cf. electronic supplement of [TUF23], Table 3).

Linear models further demonstrate that, although they assess their work as subjectively relevant, workers in essential jobs also perceive the value of their tasks to be less positive and show themselves to be less satisfied the longer they work in an occupation. In the case of employees in occupations on the "extended list", workloads are only perceived as being significantly worse in relation to physically demanding tasks and with regard to noise in the workplace (cf. [TUF23] electronic supplement, Tables 6 and 7). In contrast, work-life balance is perceived to be worse.

We also took a closer look at qualifications in essential jobs. Deviations in respect to the shares of requirement level of "no vocational qualification" and "academic qualification" exhibit differences to the totality of occupations. Employees in occupations on the "Berlin list" show a comparatively large share (21.4%) of workers stating that no vocational qualification is required to work in their jobs (all occupations: 17.5%). On the other hand, the proportion of employees with an academic requirement level (17%) lies below that for all occupations (24.4%). These ratios are reversed if we consider employees in occupations on the "extended list": The requirement level of "no vocational qualification" is significantly lower there (14.5%), whereas the proportion of "academic qualifications" is higher (27.7%).

A consideration of the qualification level also produces a similar structure. Initially, a significant preponderance of 65.23 percent of employees with vocational qualifications can be observed in occupations on the "Berlin list". In the "extended list", this share stands at 55.35 percent (all occupations: 55.67%). In occupations on the "Berlin list", 10.8% of employees have no vocational qualification. The corresponding share for the "extended list" is 7.3%. These figures are respectively above and below the value for all occupations (8.9%).

A comparison of the matching of requirements and qualification level (cf. Tables 4 and 5 in electronic supplement to [TUF23]) also shows differences between the two lists. In objective⁸ terms, employees on the "Berlin list" do not tend to be under-qualified for the

⁸ "Objective" relates to the type of measurement. A comparison is drawn between highest vocational qualification and the requirement levels according to the Classification of Occupations (KldB 2010).

work they do. Subjectively⁹, they believe they may not be deployed appropriately. By way of contrast, employees on the “extended list” are objectively more likely to be appropriately deployed rather than being under-qualified. Subjectively, however, they are more likely to work over-qualified and their deployment does not match their qualification. We were, however, unable to ascertain any statistically significant mismatch between requirements and qualification level.

Overall, our evaluations indicate a more differentiated picture of essential occupations than the one hitherto addressed in societal discourses or in academic research. Over the course of time, and in a way which diverges from the public perception of essential occupations at the beginning of the pandemic, we can assume an increase in occupations with an academic requirement level and in employees with academic training. Essential occupations, thus, have become more “average”. How this happened cannot be inferred from the Employment Survey data, which is why we turned to a different data source.

2 Reviews of essential work on twitter

Twitter is a short message service that has become established primarily in the media industry, but is also used by institutions, companies, organizations, political parties and educational institutions, as well as by private individuals. Compared to other social media there is the particularity of a limitation of text messages to a maximum of 280 characters (until November 2017 it was 140 characters). Twitter is a medium for information, communication and opinion in one place, and more or less in real time. In 2017, around 320 million people actively used the service, i.e., logged into the network, and a further 500 million followed tweets without logging in [Sa17]. Current statistics published by Twitter owner Elon Musk show 253.8 million users in 2020 [Tw20]. When the company celebrated its tenth anniversary in 2016, user figures for Germany were announced for the first time. According to these, a total of 12 million active and passive German users were recorded at the time. Although registration and use of the service is free of charge, the company behind it, Twitter Inc., is financed by advertising, evaluating user data and selling the flow of information in its network to third parties [Sa17]. Twitter uses an algorithm that is difficult for users to understand. Since the takeover of the company in October 2022 by Elon Musk, the new owner announced in march 2023 a "new era of transparency for Twitter" [Tw23], in which this should be made more comprehensible.

There is only limited information on the exact user figures of Twitter in Germany and the algorithm to which Twitter is subject. This must be considered when interpreting the data presented below and it is especially noteworthy since data in social sciences are usually either statistically representative (like the Employment Surveys analysed here) or reveal types of people, behaviours or structures (with more ethnographic, anthropological or so

⁹ “Subjective” also refers to the type of measurement. In this case, a comparison is made between highest vocational qualification and the response to the question as to which qualification is normally necessary for the exercising of one’s own occupation.

called “qualitative” data like observations or in-depth interviews). Twitter data falls somewhere between these kinds of data. On the one hand there are masses of individual data points, on the other hand they are in no way representative. Their meaning is also only revealed when looking into detailed individual, or typical, contents [Pf16].

But social media data is also a source depicting trends in opinions and discussions. We would expect topics like the pandemic to be discussed widely enough to be visible in twitter tweets of 2020 and 2021. In a sense, it is a real-time seismograph of relevant topics, where “relevant” simply stands for “important enough to many persons to write/tweet about them”. Twitter has been described very enthusiastically by young scholars as a medium which is simultaneously a mass medium as well as constructing our reality while covering all (social and individual) topics [We21]. It has as well been assessed as polarising opinions and even influencing elections [GKD19] and examined as a political sphere [BR20]. Overall, Twitter data are seen as a valuable source [Pf16].

Following our initial findings on the changes in essential jobs we want to answer one question with analysing twitter data: Is there a shift in discussions of essential jobs? The idea behind this is as follows: A change in the composition of essential jobs must have been justified. This should have happened in political discussions, as the lists are results of governmental processes. As inclusion of a job on the list of essential jobs meant rights and restrictions for their incumbents, these decisions will have been at least commented on publicly. Given the possible explanation of occupational interest groups shifting discussions and opinions towards the inclusion of specific occupations, these comments will not just have occurred after decisions had been made. A mass medium which is found to be able to foster polarisation and form opinions, even influence elections, will at least show traces of a discussion affecting people in a crisis situation like a pandemic. Whether this was the case and if we can trace such a shift in debates of essential jobs will be discussed in the following.

3 Preparation and analysis of twitter data

Since discussions as we have depicted them will include communication of sentiments, we first conducted a sentiment analysis. We focussed on tweets¹⁰ mentioning essential jobs and therefore only included tweets that had some mentioning of jobs or occupations. When such mentions were being made, they were coded with a systematic position of the German Classification of Occupations (KldB2010).¹¹ Out of all the tweets being connected to and coded with an occupation in this way we filtered out those that were on the initial “Berlin” list of essential occupations and those on the “expanded” list. It is here

¹⁰ They were scraped by students of Koblenz University, utilising BIBB’s TM4VETR (<https://github.com/TM4VETR>) and twitter’s API, checking for tweets containing mentions of official occupation names or synonyms.

¹¹ This was done automatically utilising lists of occupational names and their aliases.

where we would expect statements and comments fitting the shift in debates explained above.

After removing doubles, 1.2 million tweets remained that related to occupations on at least the expanded list from 2007 until 2023. Besides information on which list of essential jobs tweets were related to we also have information about tweet sentiment and the full text of tweets. Tweet sentiment is interesting since a change in sentiment could occur in correspondence to a shift in opinions and the debate. Tweet contents would also possibly reveal topical shifts. Due to the nature of data processing though we have tweets that contain a lot of very similar features (words, numbers, emojis and hashtagged topics). This makes topic modelling challenging, which is why we will look at predefined topics for the time being.¹²

Sentiments were assessed within the data preparation process with adapted standard procedures.¹³ Negative or positive sentiments were scarce, compared to neutral ones (four fifths neutral, around 16 percent negative, four percent positive). For our analyses we work with mean sentiments over groups of tweets (e.g. occupations or the group of essential occupations) instead of classified dichotomous measures.

While data scraping and building the dataset of tweets with sentiments, occupations and tweet texts were done with a pipeline devised by researchers at BIBB, further analysis was conducted in R (R Core Team 2023). Topics were searched for within tweets using regular expressions.

4 Assessment of essential occupations on twitter

Our data shows an increase in tweets related to occupations over time, from only 17 in 2007 up to more than 240.000 in 2023 (where scraping stopped in March) with the most tweets scraped from 2022. The share of posts relating to essential occupations from the shorter and more restrictive Berlin list show some variation, but interestingly are stable (at around 35 percent) during the pandemic, indicating that they did not raise more attention or possibly even that not more persons in these occupations tweeted about them.

Figure 1 shows the increase in absolute numbers, but only for occupations on the Berlin list. We do not see movements that could be traced back to events occurring during the pandemic, e.g. there is no peak in tweets around the start of the pandemic, only at the beginning of the first lockdown, but not at the start of the second. There does seem to be an increase around winter months, though, but we could not find seasonal shifts over time.

¹² To give an example: In topic modelling one can trim data and leave out words (features) occurring only in very few or in almost all tweets. Trimming in our data would regularly cut of more than 95 percent of features.

¹³ The routines used can be found here: <https://github.com/AliVn85/Professions-Sentiment-Analysis-on-Twitter>

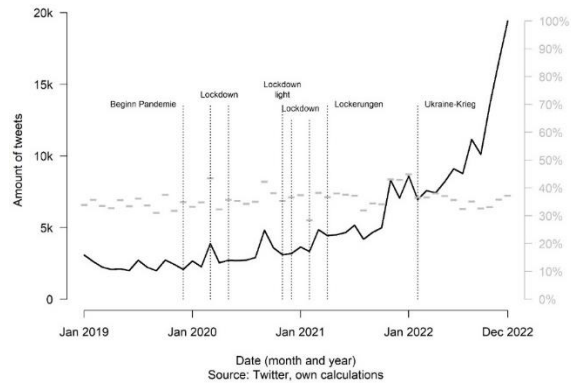


Figure 1: Tweets in essential jobs of the Berlin list

We observe negative sentiments in essential jobs to increase over time: In March 2023 we have a mean negative sentiment of .3862 in occupations on the Berlin list, in January 2019 it was .1295.

There is an interesting finding in mean sentiments regarding the differences of essential jobs of the Berlin list and the expanded list. Figure 2 shows that generally occupations of the Berlin list are connected to higher negative sentiments than those from the extended list. Taking into account that the extended list is indicative of a mean over all occupations we see an interesting pattern: Before and after the pandemic, and even within it, when there was a time of relative easement, we see that essential jobs from the Berlin list are connected to more negative sentiments than other jobs – while during the pandemic this es levelled out and at times even turned (in January, April and August 2020). This could be due to essential jobs of the Berlin list generally being discussed based on their discriminating features like long working hours, relatively low pay, relatively bad work situations. But during the pandemic there was a wave of support and appraisal towards these occupations (especially care and health jobs), which could have lead to the rapprochement of both curves. This also corresponds to workers’ self assessment (see above) of the value of their tasks as less positive in essential occupations of the Berlin list than in other occupations, as well as their work loads being higher and their work conditions being worse.

Looking at which topics were discussed figure 3 shows some interesting differences between an occupation¹⁴ from the Berlin list (sales occupations selling foodstuffs and one of the extended list (insurance and financial services). In sales occupations the negative sentiments pronounce the differences of pandemic and non-pandemic times. In the first

¹⁴ In this case we look at three-digit occupations from the German classification. These are relatively detailed regarding occupational content, as there is only one more level with the fourth digit, but they include every level of skill requirement (from no qualification needed up to academic qualifications). Even though there are many data points in our data set and we could in theory look at occupations on the five-digit level, we would run into problems of low numbers in some of these cases.

year of the pandemic, valuation was a topic with this occupation, in the second lockdown we see a peak in topics relating to acknowledging what incumbents do. In insurance and finance services we do not observe higher than average negative sentiments. Also, when they peak, they do so later than occupations from the Berlin list. Topics discussed are generally less prominent, except for a discussion of pandemic measures during the second lockdown (January 2021) and of acknowledgement right in the following month (March 2021). Valuation topics tend to be relevant during phases of easement, but never as much as with sales occupations.

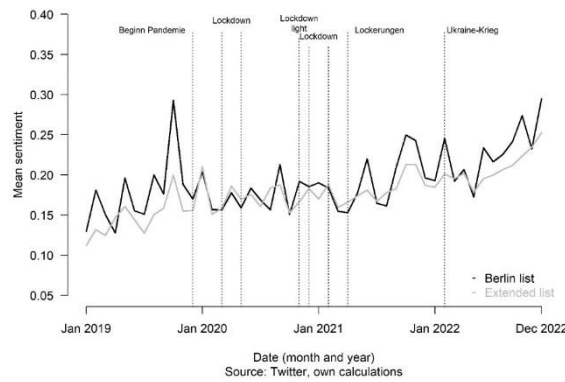


Figure 2: Negative sentiments in essential jobs

Overall the analysis of topics shows the data to reflect discussions and debates at the times they were also discussed in the media. This might be seen as a kind of verification. Combined with distinctive occupations they shed some light on how discussions on twitter were structured thematically.

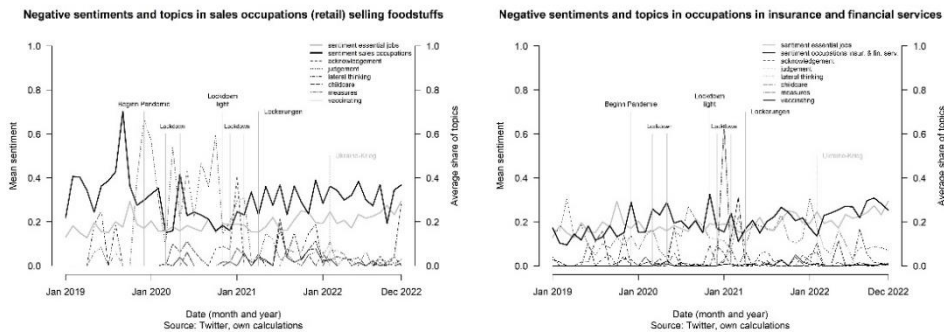


Figure 3: Topics in two essential occupations

5 Conclusion

This paper shows that valuation of essential jobs can be traced in very different data types. We see, among others, lower wages, higher mismatch with respect to their qualifications but also higher negative sentiments in occupations on the Berlin list. These findings might be seen as hints for discussions in which essential occupations were restructured in a sociological meaning. Considering developments in sentiments and topics discussed, Twitter seems to be a valuable data source. However, not least due to the ambiguous user numbers and the algorithm Twitter is subject to, the data is not easy to interpret. To further examine these initial indications, on the Twitter data in more detail and also how it can supplement other data sources, future research should establish a pipeline and examine and compare other cases, as well as other research questions in more detail.

References

- [BR20] Bouvier, G., Rosenbaum, J.E. (eds.): *Twitter, the Public Sphere, and the Chaos of Online Deliberation*. Cham, 2020.
- [GKD19] Grovera P., Kumar Kara A., Dwivedib, Y. K., Janssen, M.: Polarization and acculturation in US Election 2016 outcomes – Can twitter analytics predict changes in voting preferences? *Technological Forecasting and Social Change* 145, pp. 438-460, 2019.
- [HHR18a] Hall, A.; Hünefeld, L.; Rohrbach-Schmidt, D.: *BIBB/BAuA-Erwerbstätigenbefragung 2018 – Arbeit und Beruf im Wandel. Erwerb und Verwertung beruflicher Qualifikationen. SUF_1.0*. Bonn 2020a. doi.org/10.7803/501.18.1.1.10
- [HHR18b] Hall, A.; Hünefeld, L.; Rohrbach-Schmidt, D.: *BIBB/BAuA-Erwerbstätigenbefragung 2018 – Arbeit und Beruf im Wandel. Erwerb und Verwertung beruflicher Qualifikationen. SV_1.0*. Bonn 2020b. doi.org/10.7803/501.18.1.5.10
- [HSZ23] Herzog, L.; Sold, K.; Zimmermann, B.: *Essential Work: A Category in the Making? Unsettling the Ways We Think about Jobs, Labor, and Activities*. In: (Herzog, L.; Zimmermann, B., eds.): *Shifting Categories of Work. Unsettling the Ways We Think about Jobs, Labor, and Activities*. New York und Abingdon, pp. 252–264, 2023.
- [Ko20] Koebe, J. et.al.: *Systemrelevant, aber dennoch kaum anerkannt: Entlohnung unverzichtbarer Berufe in der Corona-Krise unterdurchschnittlich*. Berlin, 2020.
- [Pf16] Pfaffenberger, F.: *Twitter als Basis wissenschaftlicher Studien. Eine Bewertung gängiger Erhebungs- und Analysemethoden der Twitter-Forschung*. Wiesbaden, 2016.
- [R23] R Core Team: *R: A Language and Environment for Statistical Computing*. Vienna, 2023, <https://www.R-project.org>, retrieved: 13.07.2023.
- [Sa17] Samuelis, T.: *Twitter in der Bildung*. Bundeszentrale für politische Bildung. Bonn, 2017.

- [SSS21] Schrenker, A.; Samtleben, C.; Schrenker, M.: Applaus ist nicht genug. Gesellschaftliche Anerkennung systemrelevanter Berufe. Aus Politik und Zeitgeschichte 71, pp. 13–15, 2021.
- [TUF23] Tiemann, M.; Udelhofen, S.; Fournier, L.: Esteem, equivalence and valuation of essential occupations. Berufsbildung in Wissenschaft und Praxis (bwp) 52, pp. 22-26, 2023.
- [Tw20] Twitter. Elon Musk. Slides from my Twitter company talk. 2020, <https://twitter.com/elonmusk/status/1596718851097755648?s=20&t=kKOJe8nXDwbmD3i2LgXj-Q>, retrieved: 13.07.2023.
- [Tw23] Twitter. Twitter's Recommendation Algorithm. 2023, ULR: https://blog.twitter.com/engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm, retrieved: 13.07.2023.
- [We21] Wende, F.: Twitter: Das Verbreitungsmedium für alle. 140 Zeichen zur Selbstbeschreibung der Gesellschaft. Grin Verlag, 2021.