

# Occupation Coding During the Interview

## Appendix A: Additional Results

### Results for the second and third inquiry

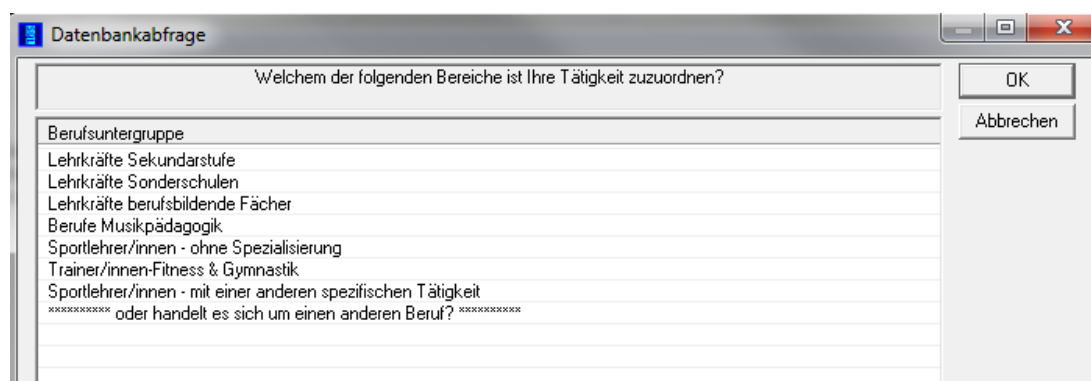


Fig. A 1: Follow-up screenshot related to figure 1 in the text. This question appears after when a respondent answers “vice principal and teacher” and selects a “different occupation” in figure 1. It shows possible 4-digit occupation sub-groups from the KIdB 2010.

Table A 1: Results for the second and third inquiry, when the respondent chooses “other occupation” in the first inquiry (cf. table 1)

Number of respondents	145	100.0%
Item nonresponse in second inquiry	8	5.5%
No occupation sub-group in second inquiry selected, break-off	77	53.1%
Occupation sub-group chosen, item nonresponse in last inquiry	1	0.7%
Occupation sub-group chosen, but no job chosen in third inquiry	28	19.3%
Occupation sub-group chosen, but selected job category from third inquiry is not in agreement with coder 1	24	16.6%
Occupation sub-group chosen and selected job category from third inquiry is in agreement with coder 1	7	4.8%

**Results in case the algorithm suggests only a single job category**

If the algorithm suggests only a single job title, not this job title but a full job description was read to the respondent, who was asked to agree, partly agree, or disagree. This applies to 30 respondents. Job descriptions were not tailored for usage in a survey and respondents often did not understand them correctly, which causes frequent errors.

Table A 2: Results in case the algorithm suggests only a single job category. This triggers a special question “Is the following description correct for your occupation?” (Question 6.23a)

	Yes	No	In parts
Manual coding by coder 1 in agreement	12	3	1
Manal coding by coder 1 in disagreement	5	6	3

**Results in case the algorithm finds more than 250 possible categories**

If more than 250 job titles are suggested by the algorithm (applicable for 10 respondents), the first list with job titles usually shown to the interviewer is skipped and occupational sub-group titles from the KIdB 2010 are asked instead. This is the same question that standard respondents get if they answer “other occupation” (Question 6.24a, Figure A1). After selecting an occupational sub-group category, respondents are provided with a detailed follow-up question in which they can select a job title (Question 6.24b).

Table A 3: Results in case the algorithm suggests at least 250 possible categories

Number of respondents	10	100%
Selected job title is in agreement with coder 1	4	40%
No job title was selected by the respondent	6	60%

## Additional answer option: similar occupation

In setting up this study, we were concerned that people would too often choose an occupation that is not the perfect choice. A small experiment was included in the survey to encourage people not to select a job title from the first inquiry but to move forward to the second inquiry regarding 4-digit occupation sub-groups, hoping that respondents are more successful to find the correct job title in the second and third enquiry. For 77 respondents, who were selected at random, the first inquiry was slightly changed and an additional answer option “or do you work in a similar occupation?” was added. When this option is selected, the interview proceeds exactly as if the answer option “different occupation” were chosen. In the rest of this paper we do not distinguish between both options and use the general term “other occupation” for both cases. Table A4 compares both experimental conditions.

Table A 4: Experiment comparison

	<i>Single answer option: “Different occupation”</i>	<i>Additional answer option: “Similar occupation”</i>
No of eligible respondents	839	76
No of respondents who select “different occupation”	139	4
No of respondents who select “similar occupation”	/	2
Proportion of eligible respondents who select “other occupation”	16.6%	7.9%
No of respondents who select a code that either agrees with coder 1 or with coder 2	520 (74% of interview-coded answers)	54 (77% of interview-coded answers)

The additional answer option was apparently not appealing and, contrary to our intention, makes it even less attractive to select “other occupation”. We calculate the odds ratio  $OR = (700/139)/(70/6) = 0.43$  and test the null hypothesis that no differences exist between one and two answer options,  $H_0: OR = 1$ . The p-value from Fisher’s Exact Test (two-sided) is significant with  $p = 0.04882$ . This result is contrary to our initial expectations and suggests that providing an additional answer option decreases the probability for respondents to select “other occupation”.

The p-value is close to 0.05 and despite its significance it is still possible that chance alone can explain it. Two arguments exist in particular, why we do not believe that there is a real difference between both conditions: (1) If it made a difference and respondents were in fact more inclined to select one of the suggested job titles whenever the additional option “similar occupation” is offered, respondents would have to choose inaccurate job titles more often and the quality would thus decrease. The last row in table A4 contrasts this logic, showing that agreement with professional coders is even higher for the respondents who get the additional answer option. (2) We also know from the analysis of interviewer behavior that, under the standard condition, the single answer option “different occupation”

was only read aloud in 18% of the interviews. It is hard to explain why this answer option is more often selected although interviewers take so little note of it.

**Can we detect from the interview if the respondent is correct?**

Respondents may have trouble to select a job title when no suggested job title is entirely correct. We hypothesized that respondents who run into difficulties answering the question are less likely to provide an accurate answer. To uncover these problematic respondents, interviewers were asked if the respondent found it difficult to answer the question (no. 6.26 in appendix B). Table A5 contains the results. Short hesitation appears to be an indication of erroneous interview coding. However, the difference to standard behavior is not pronounced enough to use this characteristic for discrimination between accurate and questionable codes.

Table A 5: Analysis of respondent’s behavior in interview coding

<i>Behavior</i>	<i>Freq. Behavior</i>	<i>Freq. Correct</i>	<i>Proportion Correct</i>
No anomalies	657	496	75.5%
Short hesitation	86	56	65.1%
Thinking for seconds	15	12	80.0%
Thinking aloud resp. asking queries	11	9	81.8%
Not applicable	1	1	100.0%

The second column shows frequencies for various behaviors. The third column shows how often the interview-coded category is in agreement with at least one professional coder. The last column gives the proportion how often the interview-coded category is in agreement, given the specified behaviour (i.e., Freq. Correct / Freq Behavior).