

**Title: German value set for the EQ-5D-5L**

**Journal:** Pharmacoeconomics

**Authors:** Kristina Ludwig, J-Matthias Graf von der Schulenburg, Wolfgang Greiner

**Corresponding author:** Kristina Ludwig, Health Economics and Health Care Management, Bielefeld University, E-mail: kristina.ludwig@uni-bielefeld.de

**Supplementary material 1: Final QC report**

QC Report for EQVT study – Germany

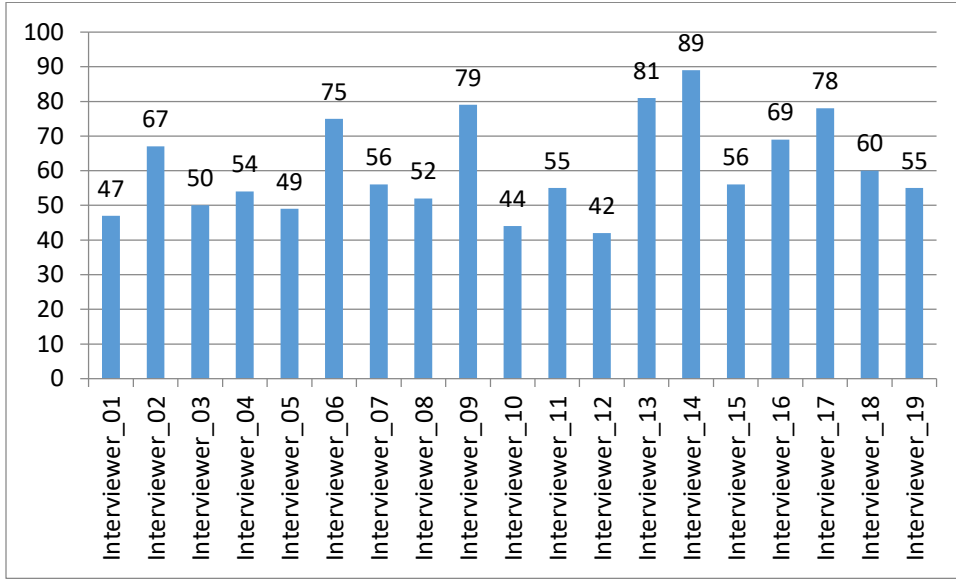
Total number of interviews: 1158

**Table 1: Sample demographics**

<b>Age</b>	<b>FEMALE</b>	<b>MALE</b>	<b>Total</b>	<b>%</b>
<b>&lt;25</b>	46	48	<b>94</b>	8.12%
<b>[25 - 34]</b>	72	64	<b>136</b>	11.74%
<b>[35 - 44]</b>	103	88	<b>191</b>	16.49%
<b>[45 - 54]</b>	124	110	<b>234</b>	20.21%
<b>[55 - 64]</b>	115	98	<b>213</b>	18.39%
<b>[65 - 74]</b>	84	80	<b>164</b>	14.16%
<b>&gt;75</b>	74	52	<b>126</b>	10.88%
<b>Total</b>	<b>618</b>	<b>540</b>	<b>1158</b>	<b>100.00%</b>

This table shows the total number of respondents in each age-sex category.

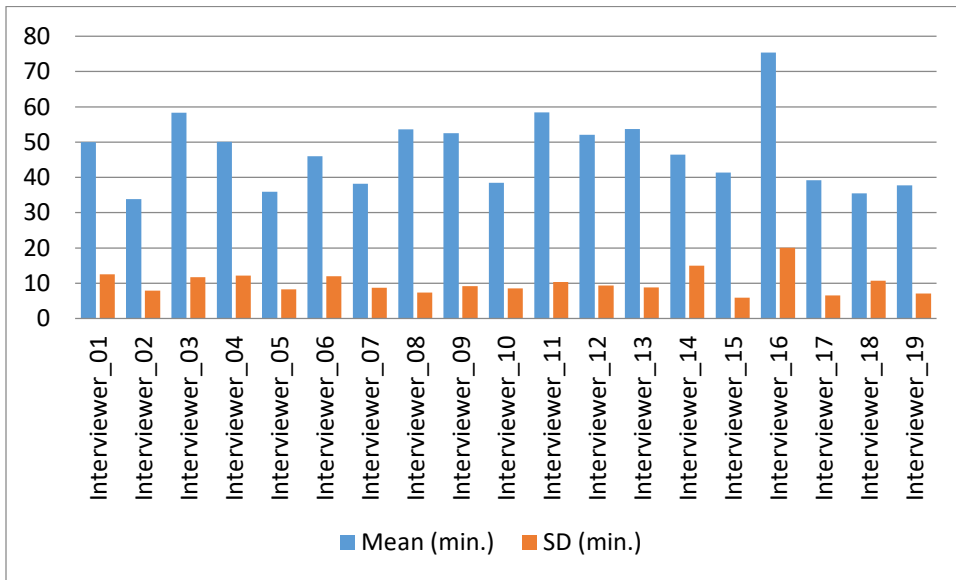
**Figure 1. Number of interviews completed, by interviewer**



This figure shows the total number of interviews completed by each interviewer.

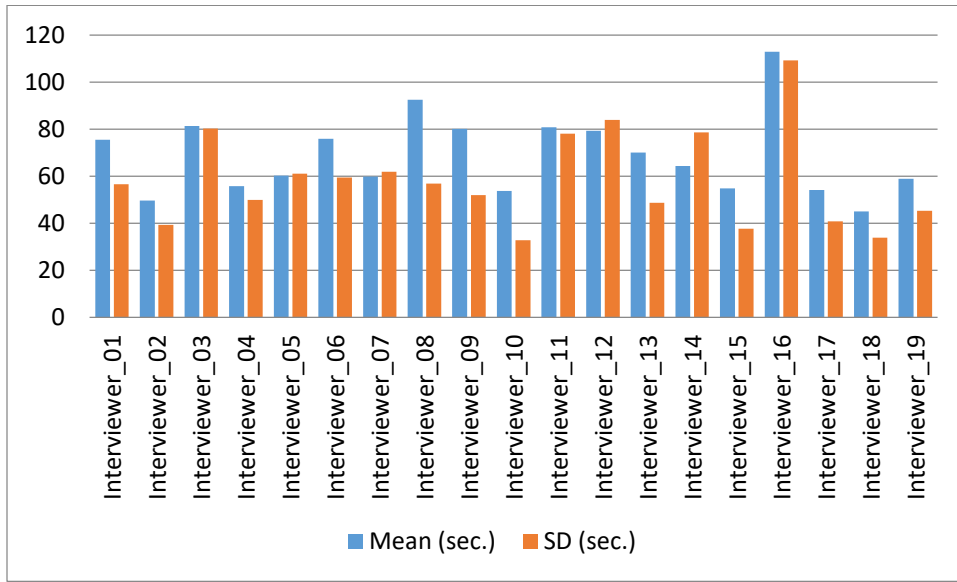
## Protocol compliance, by interviewer

**Figure 2. Duration of interviews, by interviewer**



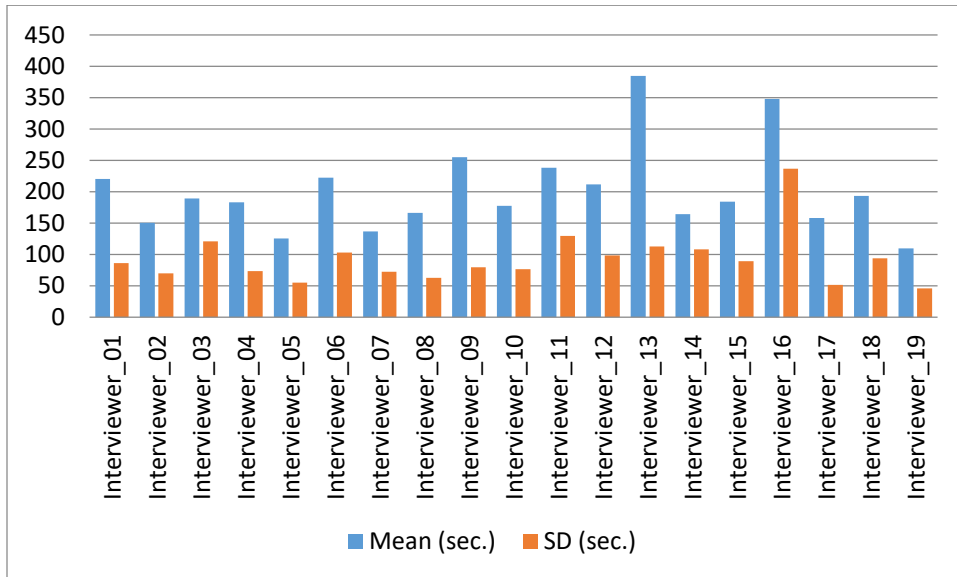
This figure shows the mean (and standard deviation) amount of time taken (in minutes) to complete the valuation questionnaire, by interviewer. This excludes any time taken to complete additional questionnaires such as the country-specific background questionnaire.

**Figure 3. Time taken to complete a single TTO task, by interviewer**



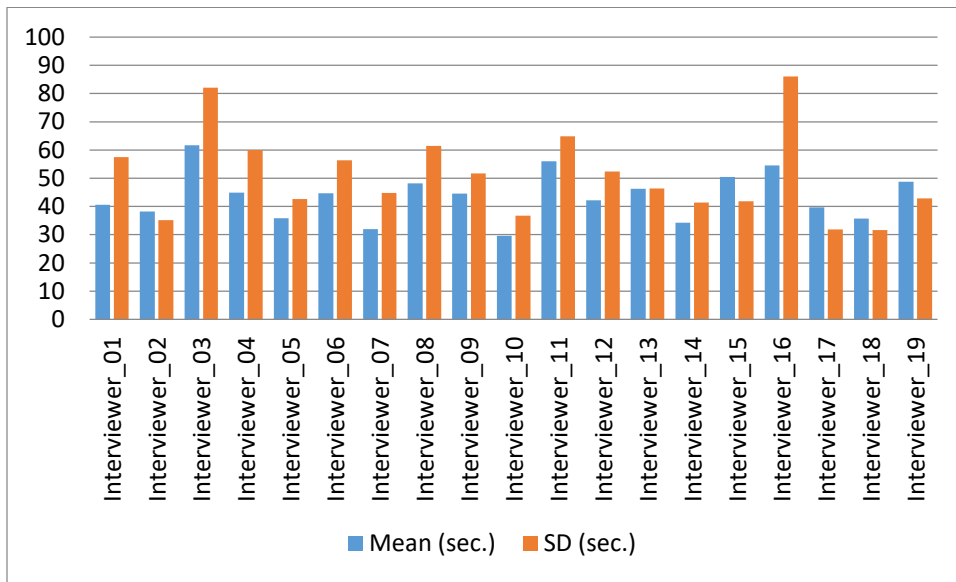
This figure shows the mean (and standard deviation) amount of time taken (in seconds) to complete each TTO task, by interviewer. This excludes the wheelchair example and practice TTO tasks.

**Figure 4. Time spent on feedback module, by interviewer**



This figure shows the mean (and standard deviation) amount of time taken (in seconds) to complete the feedback module, by interviewer.

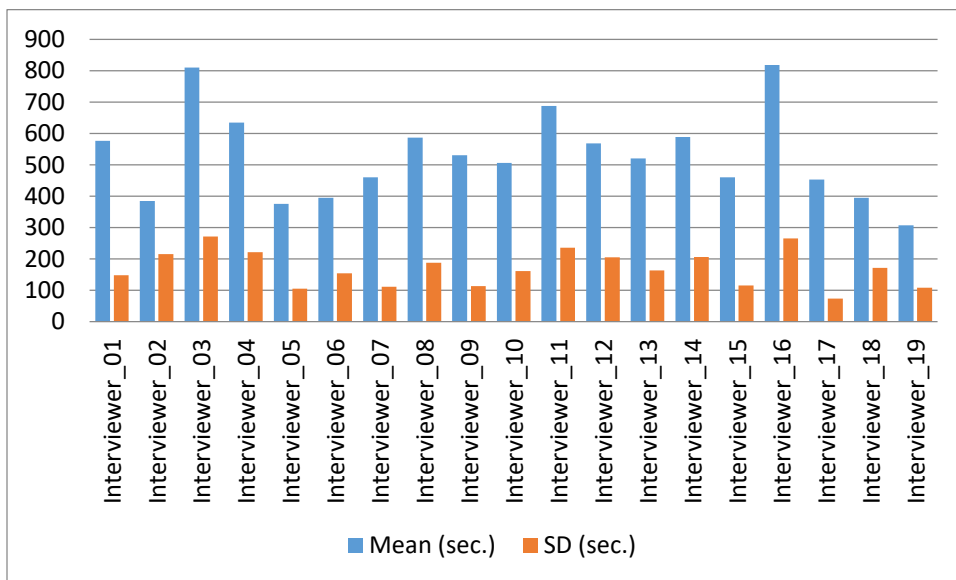
**Figure 5. Time taken to complete a single DC task, by interviewer**



This figure shows the mean (and standard deviation) amount of time taken (in seconds) to complete each DC task, by interviewer.

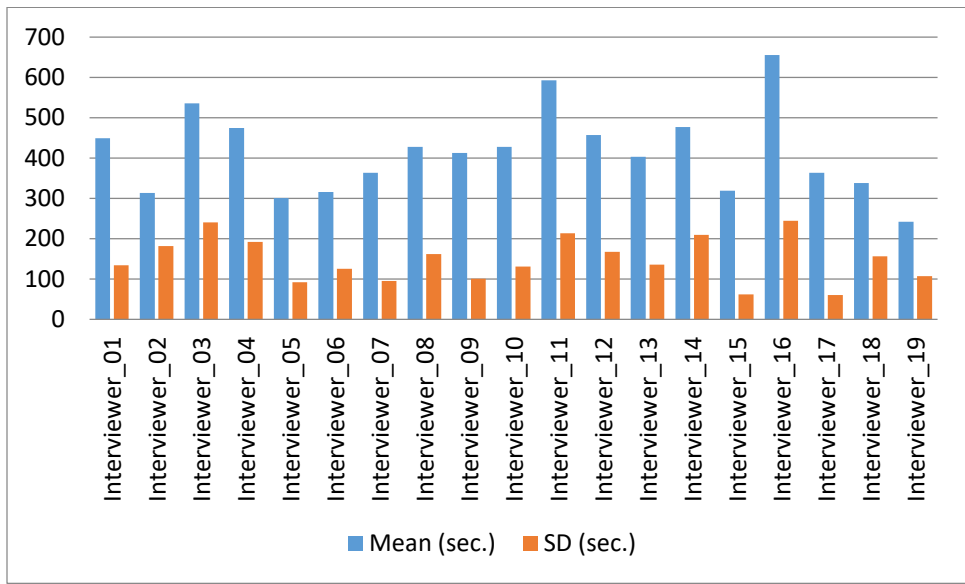
## Wheelchair example stats.

**Figure 6. Time spent on TTO wheelchair example, by interviewer**



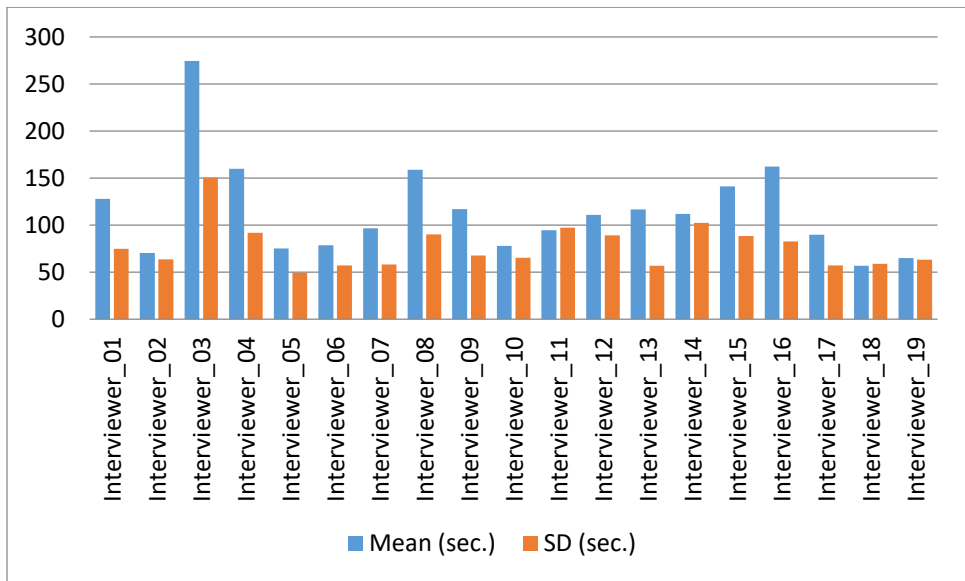
This figure shows the mean (and standard deviation) amount of time spent (in seconds) on the wheelchair example designed to introduce the TTO task, by interviewer.

**Figure 7. Time spent on BTD element of TTO wheelchair example, by interviewer**



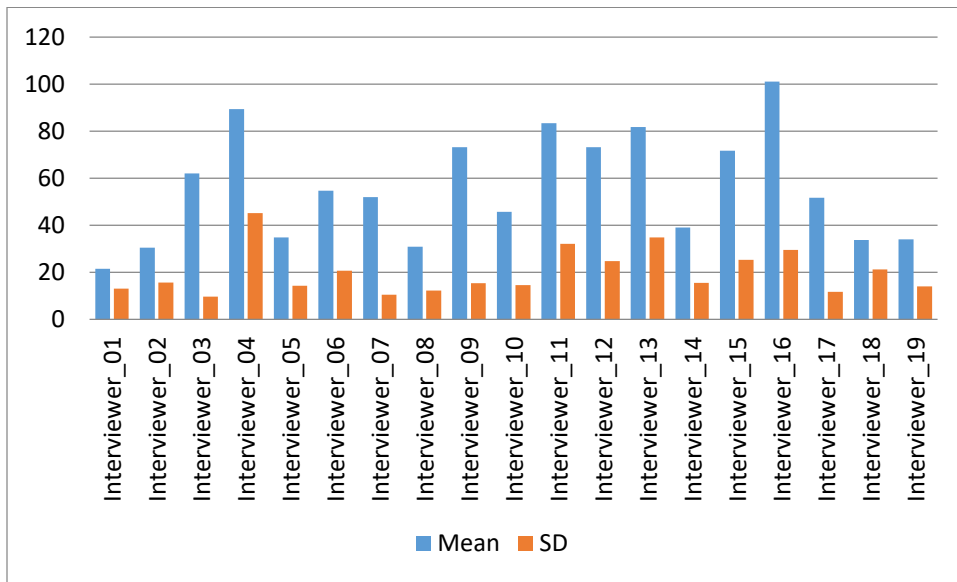
This figure shows the mean (and standard deviation) amount of time spent (in seconds) on the better-than-dead element of the wheelchair example, by interviewer.

**Figure 8. Time spent on WTD element of TTO wheelchair example, by interviewer**



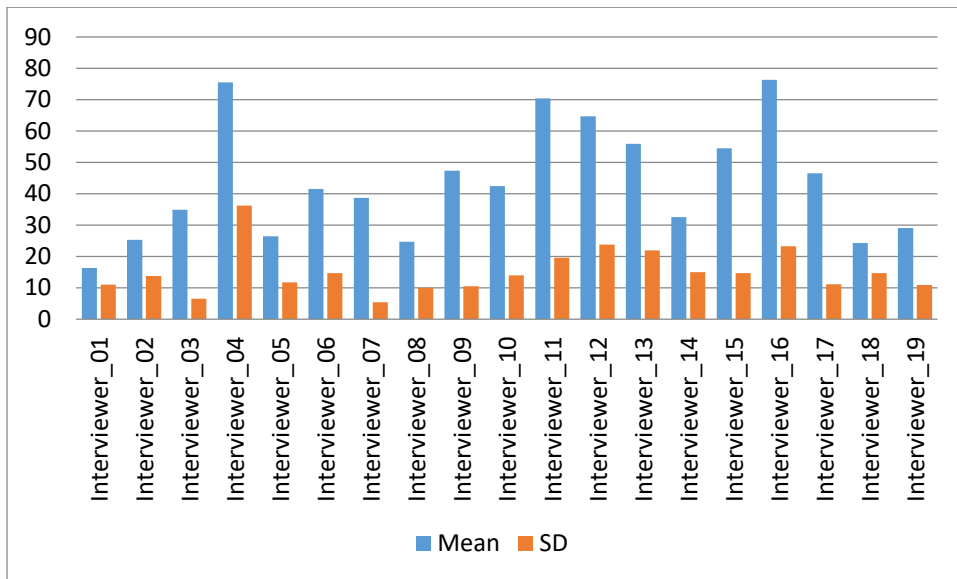
This figure shows the mean (and standard deviation) amount of time spent (in seconds) on the worse-than-dead element of the wheelchair example (designed to introduce the lead time TTO task), by interviewer.

**Figure 9. Number of moves used to complete TTO wheelchair example, by interviewer**



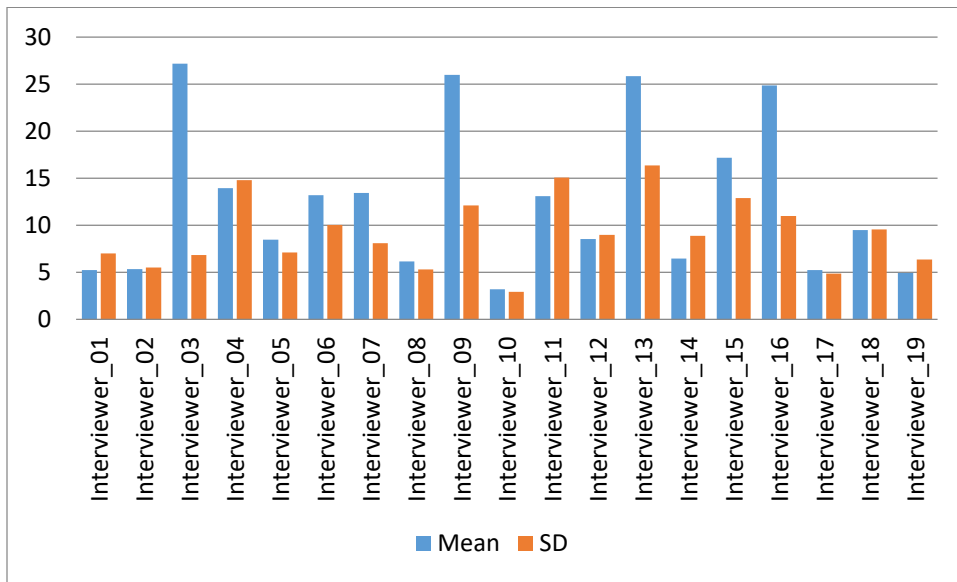
This figure shows the mean (and standard deviation) number of iterative steps used in the wheelchair example, by interviewer.

**Figure 10. Number of moves used in BTD element of TTO wheelchair example, by interviewer**



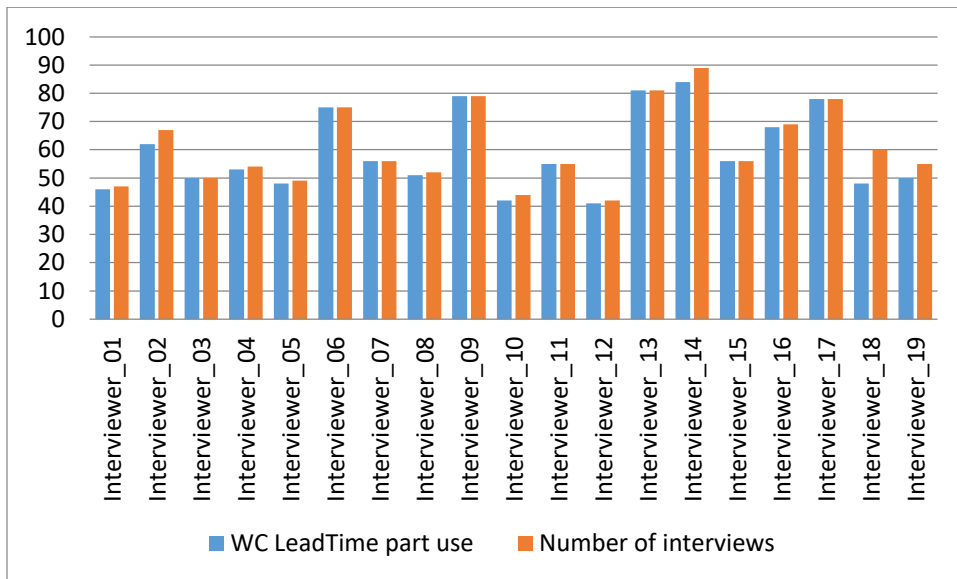
This figure shows the mean (and standard deviation) number of iterative steps used in the better-than-dead element of the wheelchair example, by interviewer.

**Figure 11. Number of moves used in WTD element of TTO wheelchair example, by interviewer**



This figure shows the mean (and standard deviation) number of iterative steps used in the worse-than-dead element of the wheelchair example, by interviewer.

**Figure 12. Use of WTD element of TTO wheelchair example, by interviewer**



This figure shows the number of interviews in which the worse-than-dead element of the wheelchair example was used, by interviewer. The total number of interviews completed by each interviewer is also shown in this figure, for comparison purposes.

**Table 2. Flagged interviews**

Interviewer	N	N flagged	% flagged	WC LT	% WC LT	Incon size	% Incon size	WC time	% WC time	TTO time	% TTO time
Interviewer_01	47	2	4%	1	2%	1	2%	0	0%	0	0%
Interviewer_02	67	17	25%	5	7%	0	0%	13	19%	9	13%
Interviewer_03	50	0	0%	0	0%	0	0%	0	0%	0	0%
Interviewer_04	54	8	15%	1	2%	1	2%	1	2%	6	11%
Interviewer_05	49	4	8%	1	2%	0	0%	0	0%	3	6%
Interviewer_06	75	0	0%	0	0%	0	0%	0	0%	0	0%
Interviewer_07	56	3	5%	0	0%	1	2%	0	0%	2	4%
Interviewer_08	52	1	2%	1	2%	0	0%	0	0%	0	0%
Interviewer_09	79	0	0%	0	0%	0	0%	0	0%	0	0%
Interviewer_10	44	4	9%	2	5%	1	2%	1	2%	1	2%
Interviewer_11	55	0	0%	0	0%	0	0%	0	0%	0	0%
Interviewer_12	42	1	2%	1	2%	0	0%	1	2%	1	2%
Interviewer_13	81	1	1%	0	0%	0	0%	1	1%	1	1%
Interviewer_14	89	8	9%	5	6%	1	1%	1	1%	3	3%
Interviewer_15	56	0	0%	0	0%	0	0%	0	0%	0	0%
Interviewer_16	69	2	3%	1	1%	0	0%	2	3%	1	1%
Interviewer_17	78	1	1%	0	0%	0	0%	0	0%	1	1%
Interviewer_18	60	19	32%	12	20%	5	8%	1	2%	9	15%
Interviewer_19	55	12	22%	5	9%	1	2%	6	11%	1	2%

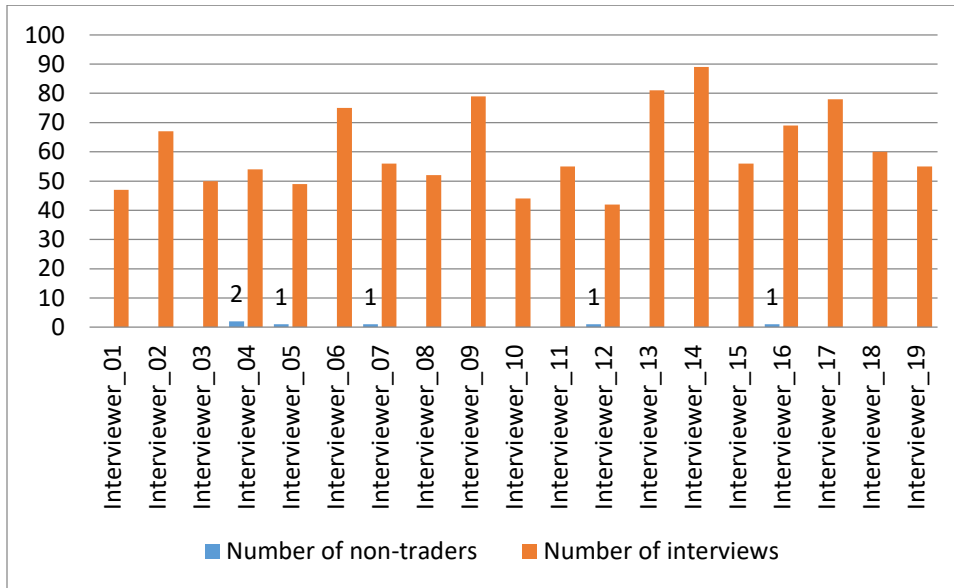
This table shows how many times each interviewer's TTO data have been flagged for data quality reasons. The total number of flagged interviews is shown in column 2, and the proportion of flagged interviews is shown in column 3. A given interview may be flagged for more than one reason. The flags are defined as follows:

- 1) WC LT - Interview is flagged if the interviewer does not enter the worse-than-dead element of the wheelchair example
- 2) Incon size - Interview is flagged if the respondent has a clear inconsistency in their TTO ratings (the value for 55555 is not the lowest and is at least 0.5 higher than that of the state with the lowest value).
- 3) WC time - Interview is flagged if the interviewer does not spend at least 180 seconds (3 minutes) on the wheelchair example.
- 4) TTO time - Interview is flagged if the respondent does not spend at least 5 minutes on the 10 TTO tasks



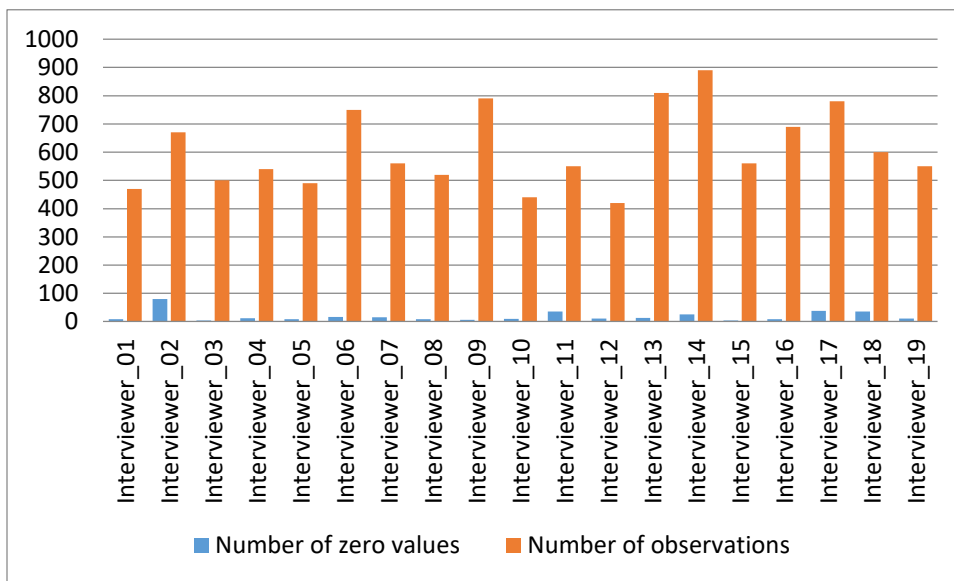
## Face validity of the data (before feedback module), by interviewer

Figure 13. Non-traders, by interviewer



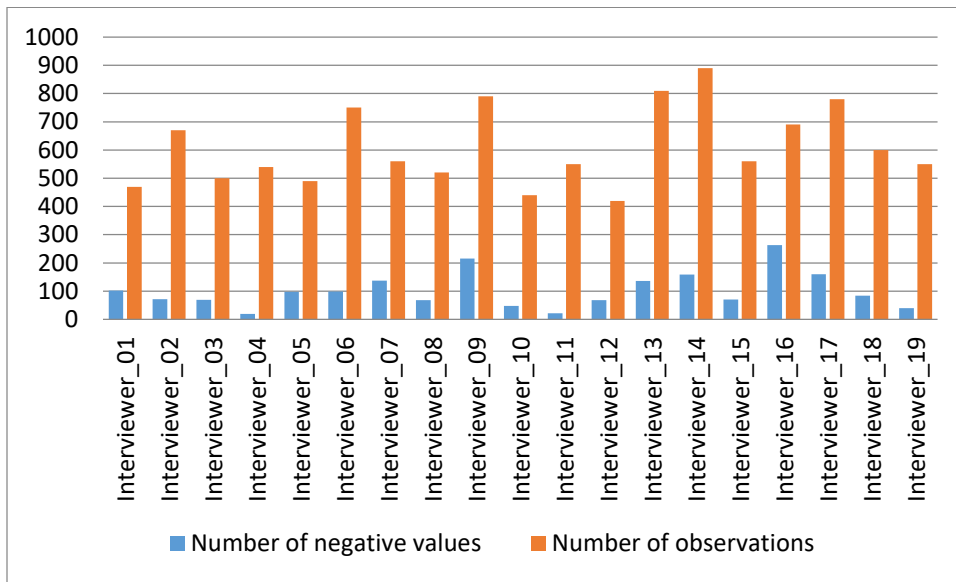
This figure shows, by interviewer, the total number of respondents who can be described as a 'non-trader' - that is, a respondent who failed to trade in any of the TTO tasks and therefore gave a value of 1 for all health states. The total number of interviews completed by each interviewer is also shown in this figure, for comparison purposes.

Figure 14. Zero values, by interviewer



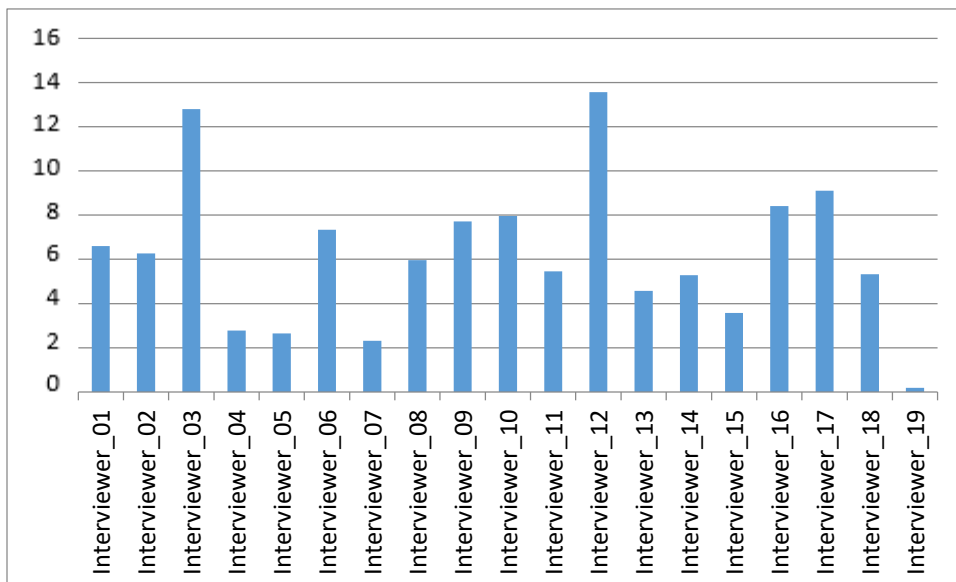
This figure shows, by interviewer, the total number of health states given a value of exactly 0 in the TTO tasks. The total number of observations for each interviewer (total number of interviews multiplied by 10) is also shown in this figure, for comparison purposes.

**Figure 15. Negative values, by interviewer**



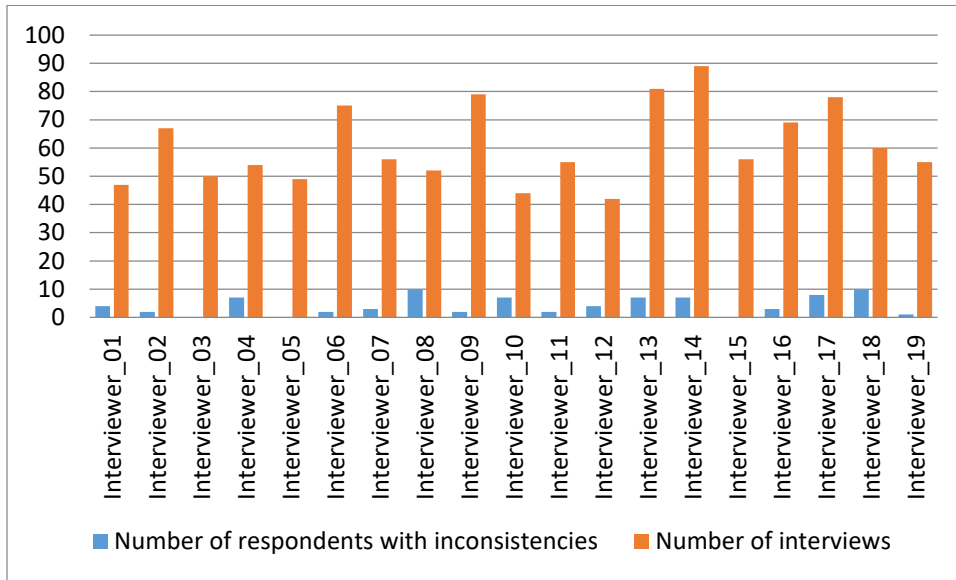
This figure shows, by interviewer, the total number of health states given a value of less than 0 in the TTO tasks (i.e. the number of times a health state was valued as being worse than dead). The total number of observations for each interviewer (total number of interviews multiplied by 10) is also shown in this figure, for comparison purposes.

**Figure 16. Percentage of health states flagged using feedback module**



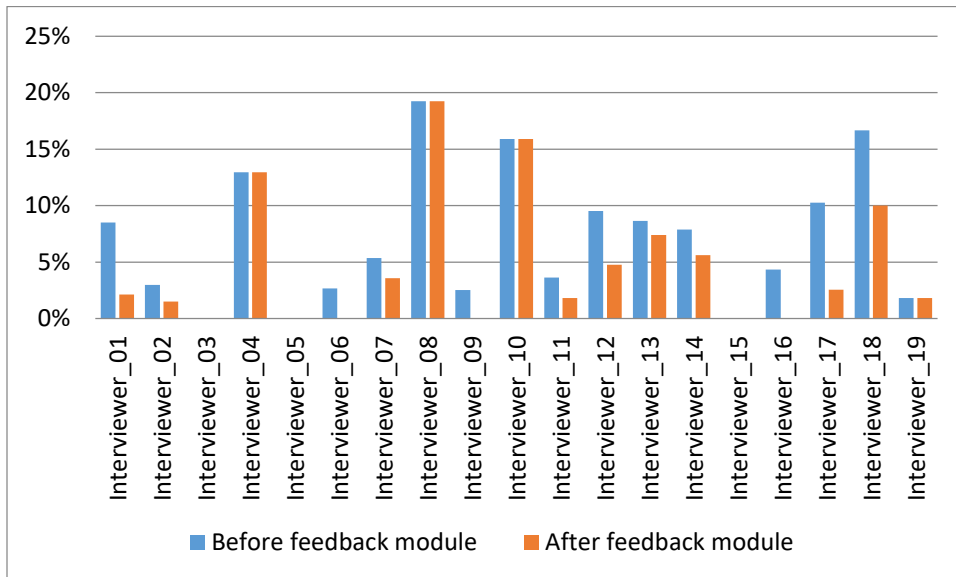
This figure shows, by interviewer, the overall proportion of health states that were flagged for exclusion by respondents via the feedback module.

**Figure 17. Respondents with inconsistencies involving 55555, by interviewer**



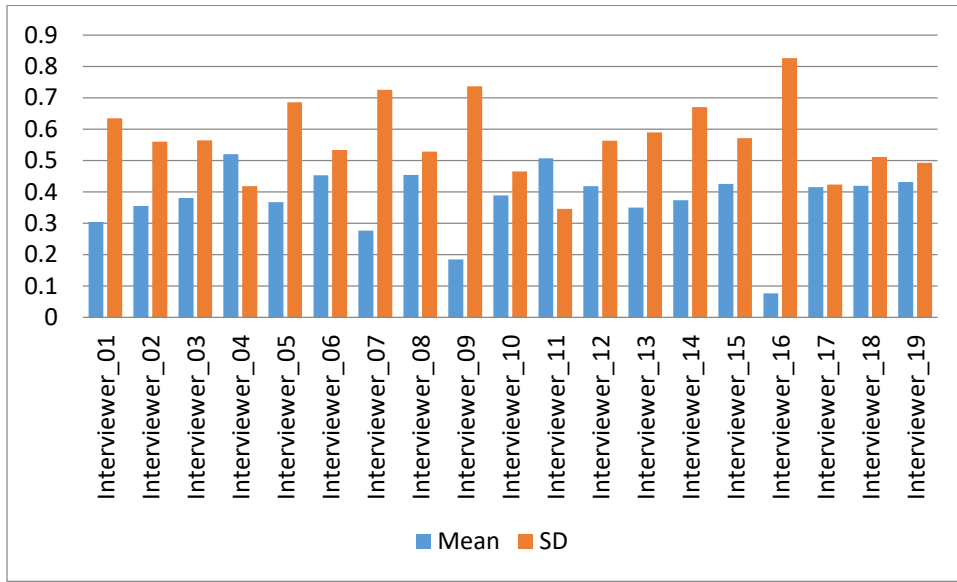
This figure shows, by interviewer, the number of respondents whose TTO data contain at least one 'inconsistency' in relation to health state 55555. An inconsistency is defined as a case where 55555 was not given the lowest value. The total number of interviews completed by each interviewer is also shown in this figure, for comparison purposes.

**Figure 18. Proportion of respondents with inconsistencies before and after the feedback module, by interviewer**



This figure shows, by interviewer, the proportion of respondents whose TTO data contained at least one inconsistency (as defined above) both before and after they were given the opportunity to flag data for removal using the feedback module.

Figure 19. Mean TTO value before feedback module, by interviewer



This figure shows the mean (and standard deviation) value observed across all TTO tasks, by interviewer. This excludes the wheelchair example and practice TTO tasks.

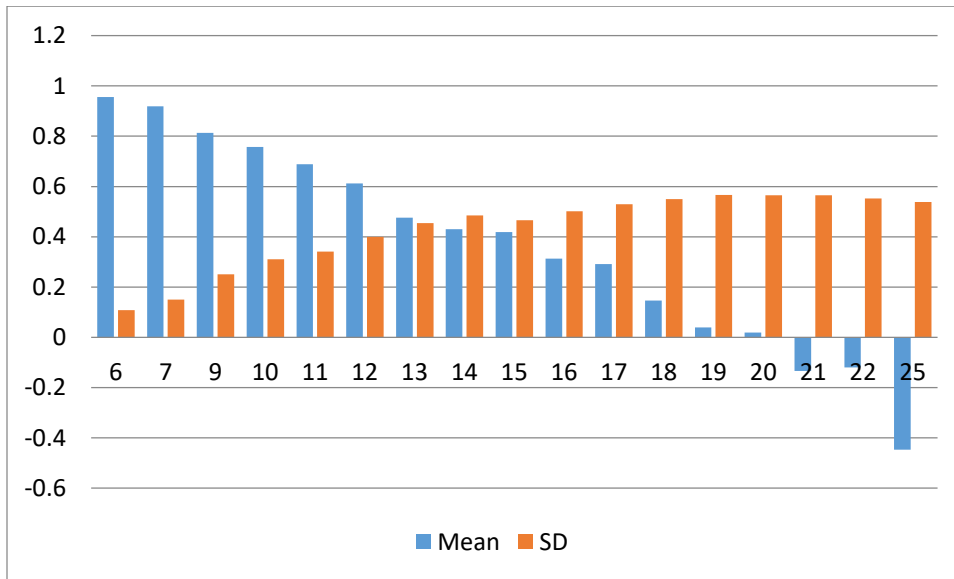
**Table 3. DCE unusual responses**

Interviewer	N	Time (min.)	IF AAAAAAA	IF BBBBBBB	IF ABABABA	IF BABABAB
Interviewer_01	47	8.79	0	0	0	0
Interviewer_02	67	8.28	1	0	1	0
Interviewer_03	50	13.35	0	0	1	1
Interviewer_04	54	9.74	0	2	0	0
Interviewer_05	49	7.76	0	0	0	0
Interviewer_06	75	9.69	0	0	0	1
Interviewer_07	56	6.92	1	0	1	0
Interviewer_08	52	10.43	0	0	0	0
Interviewer_09	79	9.65	1	0	1	1
Interviewer_10	44	6.42	0	0	1	1
Interviewer_11	55	12.13	0	0	2	0
Interviewer_12	42	9.15	0	0	0	0
Interviewer_13	81	10.03	0	0	1	0
Interviewer_14	89	7.42	0	0	1	0
Interviewer_15	56	10.93	0	0	1	0
Interviewer_16	69	11.81	0	1	0	2
Interviewer_17	78	8.60	1	0	1	1
Interviewer_18	60	7.74	0	0	0	0
Interviewer_19	55	10.56	0	0	1	0

This table shows, by interviewer: the number of interviews completed (column 2); the mean amount of time taken (in minutes) to complete the 7 DC tasks (column 3); and the number of respondents who gave unusual sets of choices across all seven DCE tasks (columns 4-7). For example, if the respondent chose state A in all seven tasks, this is flagged in column 4.

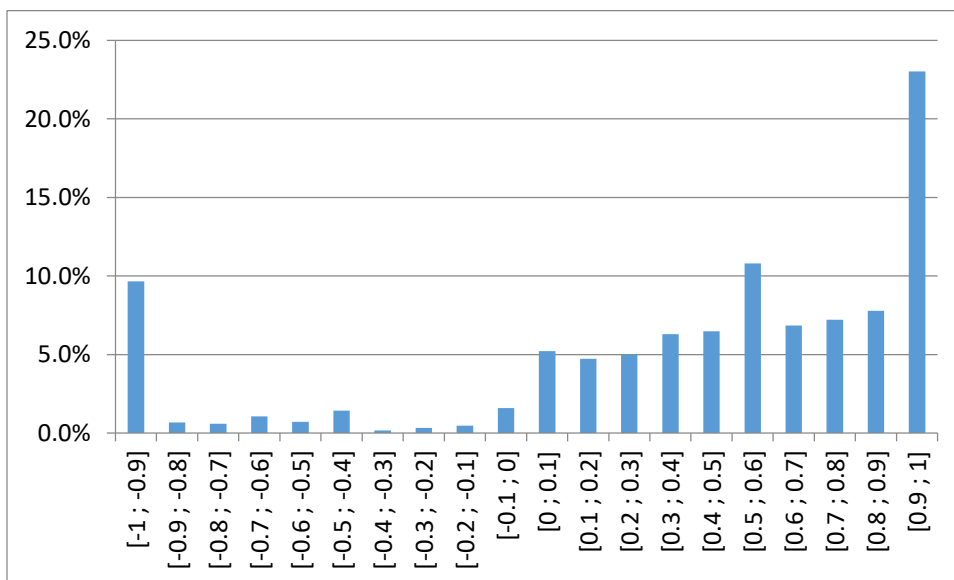
## Face validity of aggregate data.

Figure 20. Mean TTO value before feedback module, by level sum score



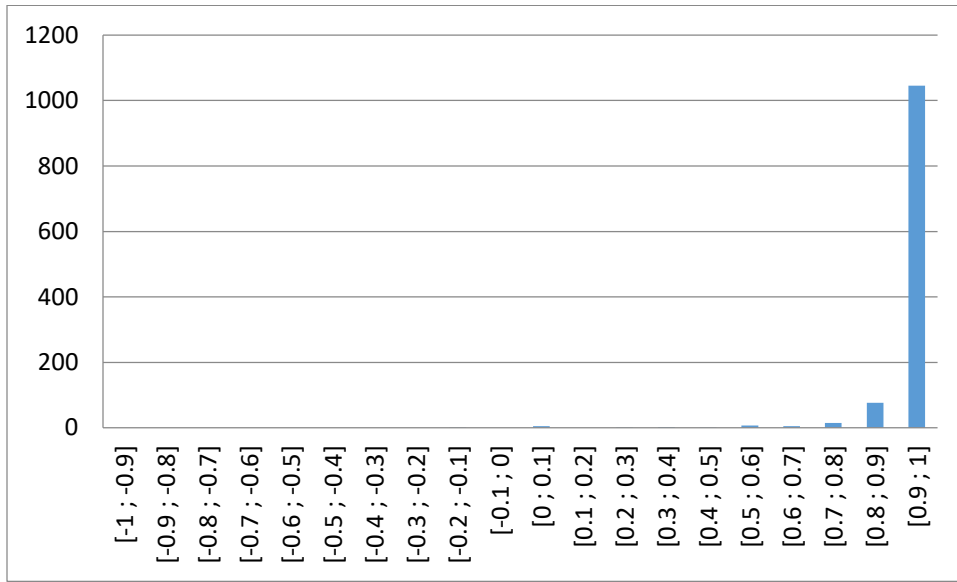
This figure shows the mean (and standard deviation) TTO value observed, by level sum score, across all interviewers. The level sum score is a proxy for severity and is calculated by summing the five dimension levels for each health state. We would expect health states with lower level sum scores (e.g. 21111: 2+1+1+1+1=6) to have higher mean values than those with higher level sum scores (e.g. 55555: 5+5+5+5+5=25). This excludes the wheelchair example and practice TTO tasks.

Figure 21. Overall TTO value distribution before feedback module



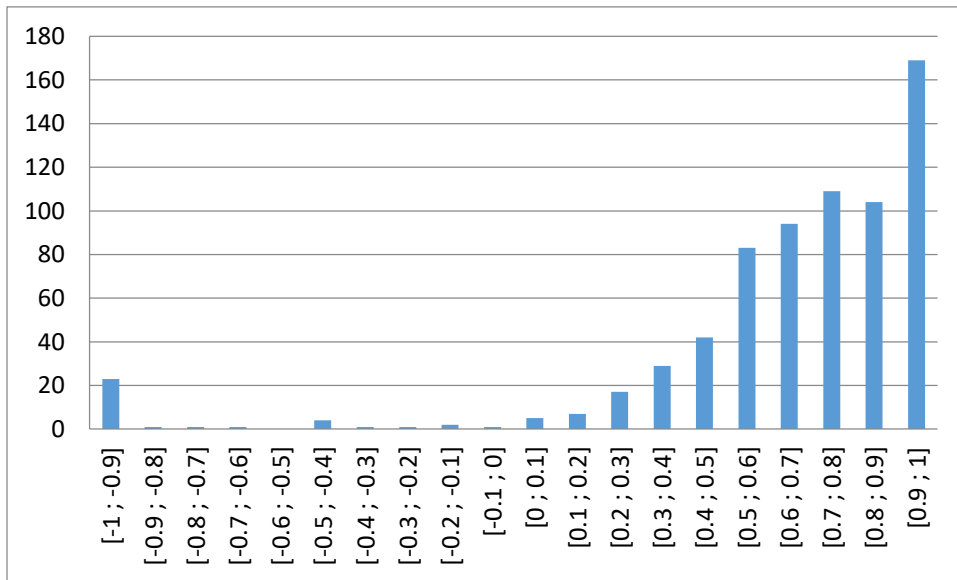
This figure shows the TTO value distribution for all health states. For example, the rightmost bar shows the proportion of observations of values greater than 0.9 and less than or equal to 1.0. This excludes the wheelchair example and practice TTO tasks.

**Figure 22. TTO value distribution before feedback module: level sum score = 6**



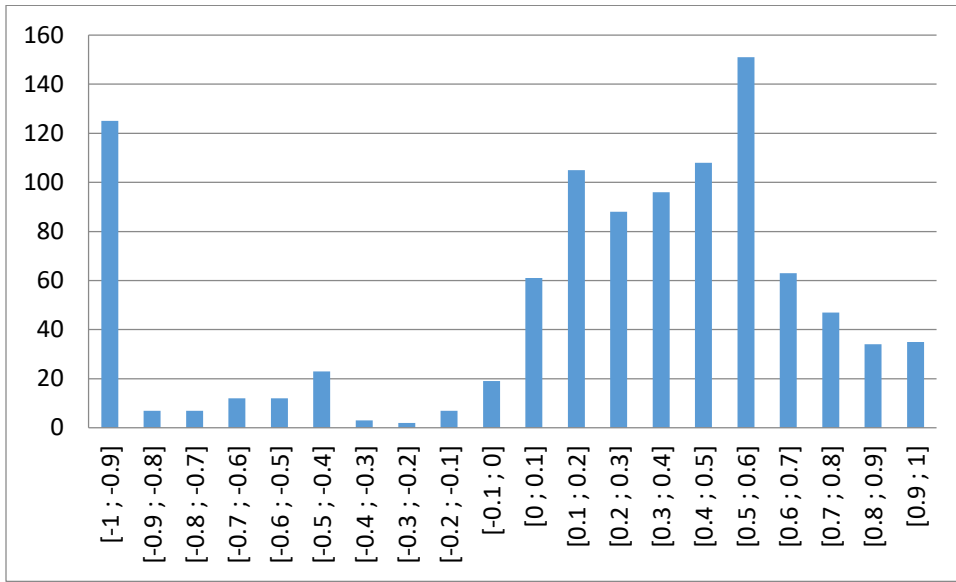
This figure shows the TTO value distribution for health states with a level sum score of 6 (e.g. 21111). This excludes the wheelchair example and practice TTO tasks.

**Figure 23. TTO value distribution before feedback module: level sum score = 12**



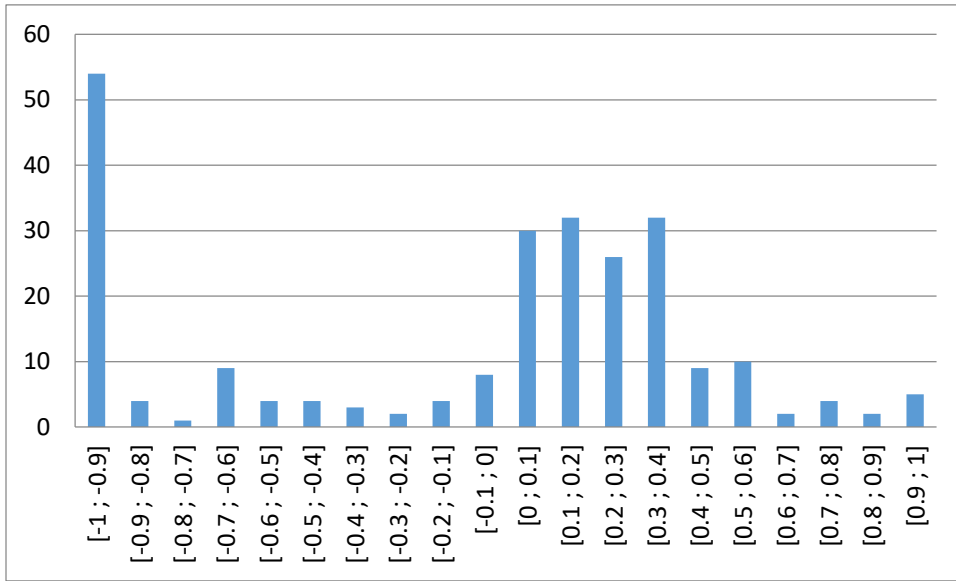
This figure shows the TTO value distribution for health states with a level sum score of 12 (e.g. 52221). This excludes the wheelchair example and practice TTO tasks.

**Figure 24. TTO value distribution before feedback module: level sum score = 18**



This figure shows the TTO value distribution for health states with a level sum score of 18 (e.g. 54342). This excludes the wheelchair example and practice TTO tasks.

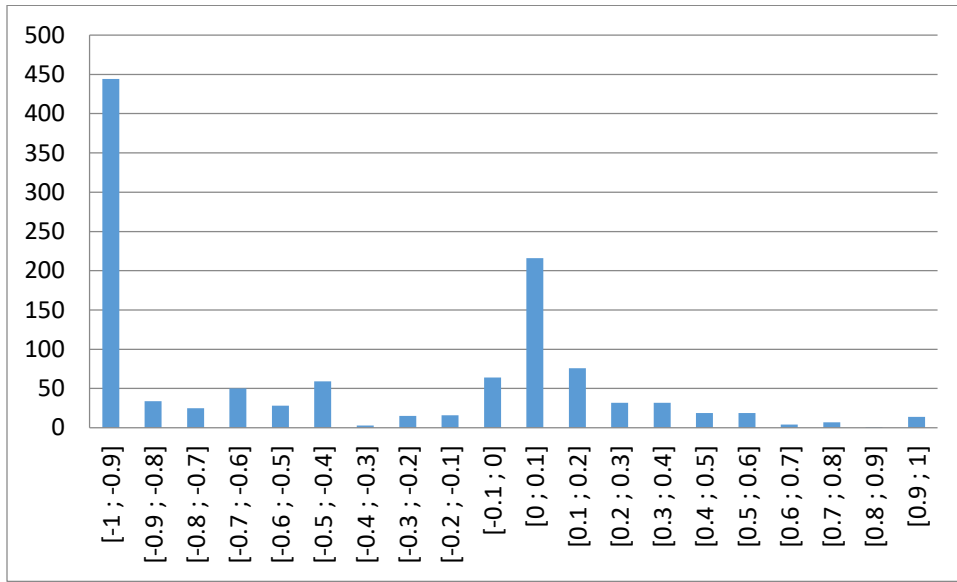
**Figure 25. TTO value distribution before feedback module: level sum score = 21**



This figure shows the TTO value distribution for health states with a level sum score of 21 (e.g. 44553). This excludes the wheelchair example and practice TTO tasks.



**Figure 26. TTO value distribution before feedback module: level sum score = 25**



This figure shows the TTO value distribution for health states with a level sum score of 25 ((the worst health state in the descriptive system - 55555)). This excludes the wheelchair example and practice TTO tasks.