

Contents

1	Test Information	2
2	Student Progress	2
3	Test Statistics	6
3.1	Summary Statistics	6
3.2	Grade Boundaries	8
3.3	Response Summary	9
3.3.1	PT12 Data	9
3.3.2	Longitudinal Data	11

List of Tables

1	Student Progress	2
2	Student Progress (continued)	3
3	Year 2 Descriptive Statistics	6
4	Year 3 Descriptive Statistics	6
5	Year 4 Descriptive Statistics	6
6	Year 5 Descriptive Statistics	6
7	Year 6 Descriptive Statistics	7
8	Distribution of students per cohort that will be awarded different PT grades	8
9	Cut Scores	8
10	Mean Number of Responses to 120 Items	9
11	Percentage of Responses	11

List of Figures

1	Longitudinal Mean Progress Test Scores	4
2	Longitudinal Progress Test Scores (Mean and 95% CI)	5
3	Boxplots of PT12 Scores By Cohort	7
4	Response Summary Boxplots	9
5	Response Summary Boxplots by Cohort and PT12 Grade	10
6	Summaries of Responses	12

1 Test Information

This document contains the results of the **FINAL** analysis run of the University of Auckland MBChB Progress Test 12 (PT12) data and it was specifically prepared for the students who sat the said test.

The following items were excluded in this run : Q016, Q029, Q063, Q078, Q117

Other Information.

2 Student Progress

Table 1 and Table 2 show the mean progress test scores of each cohort from the first to the latest test that they sat. The *Cohort* column shows the year that most of the students in that group entered the programme. The *Year* column on the other hand tells the year level of the cohort when they sat a particular progress test.

Table 1: Student Progress

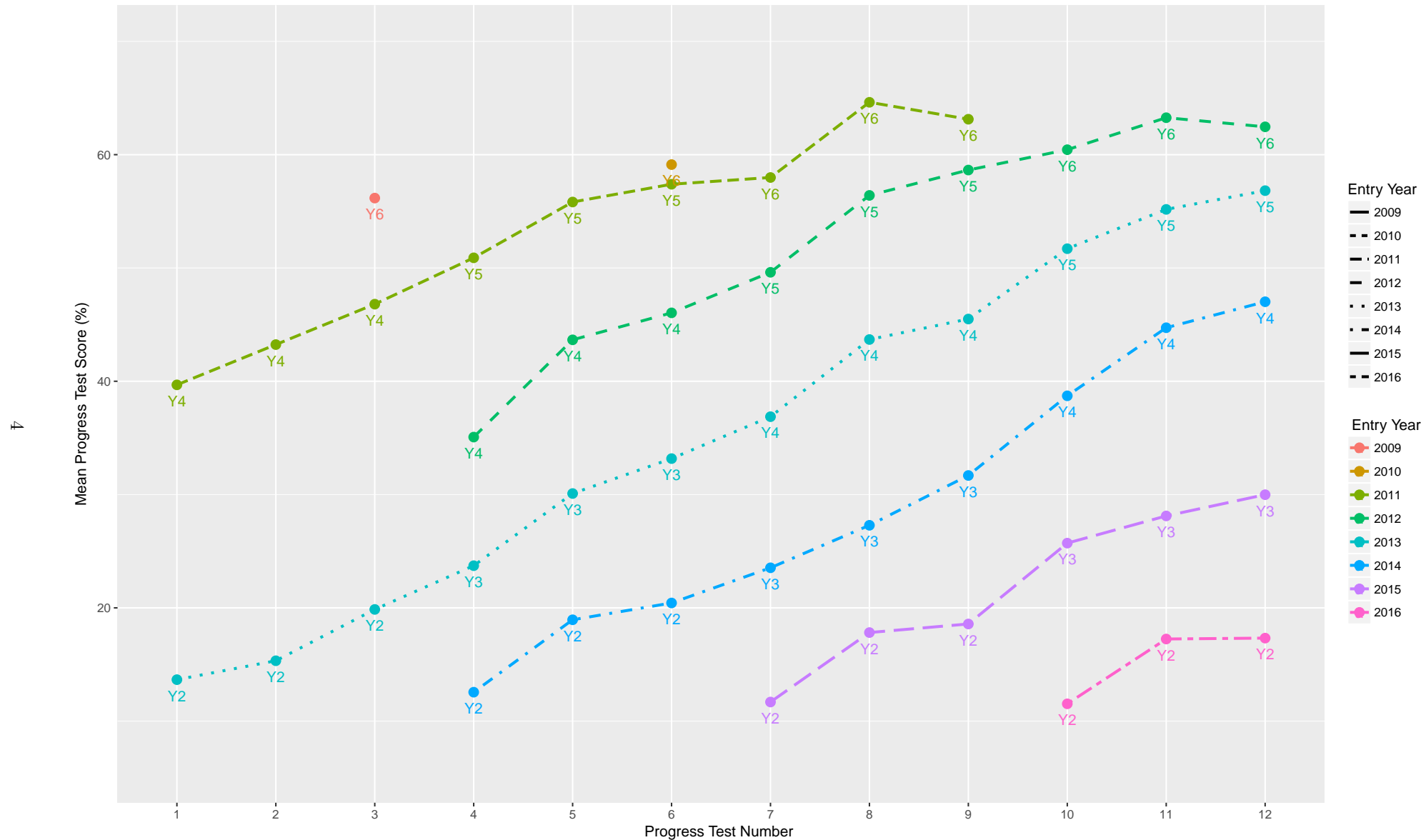
Cohort	PT Number	Year	Sitting	Mean PT Score
2013	1	2	241	13.7
2011	1	4	203	39.7
2013	2	2	238	15.3
2011	2	4	201	43.2
2013	3	2	240	19.9
2011	3	4	202	46.8
2009	3	6	123	56.2
2014	4	2	260	12.6
2013	4	3	235	23.7
2012	4	4	236	35.1
2011	4	5	199	50.9
2014	5	2	257	18.9
2013	5	3	233	30.1
2012	5	4	235	43.7
2011	5	5	198	55.8
2014	6	2	257	20.4
2013	6	3	234	33.2
2012	6	4	231	46.0
2011	6	5	195	57.4
2010	6	6	138	59.1

Table 2: Student Progress (continued)

Cohort	PT Number	Year	Sitting	Mean PT Score
2015	7	2	283	11.7
2014	7	3	255	23.5
2013	7	4	234	36.9
2012	7	5	235	49.6
2011	7	6	156	58.0
2015	8	2	280	17.8
2014	8	3	253	27.3
2013	8	4	232	43.7
2012	8	5	234	56.4
2011	8	6	149	64.6
2015	9	2	280	18.6
2014	9	3	253	31.7
2013	9	4	232	45.5
2012	9	5	234	58.6
2011	9	6	194	63.1
2016	10	2	279	11.5
2015	10	3	279	25.7
2014	10	4	246	38.7
2013	10	5	226	51.7
2012	10	6	177	60.4
2016	11	2	279	17.3
2015	11	3	279	28.1
2014	11	4	246	44.7
2013	11	5	221	55.2
2012	11	6	164	63.3
2016	12	2	277	17.3
2015	12	3	277	30.0
2014	12	4	246	47.0
2013	12	5	218	56.8
2012	12	6	214	62.5

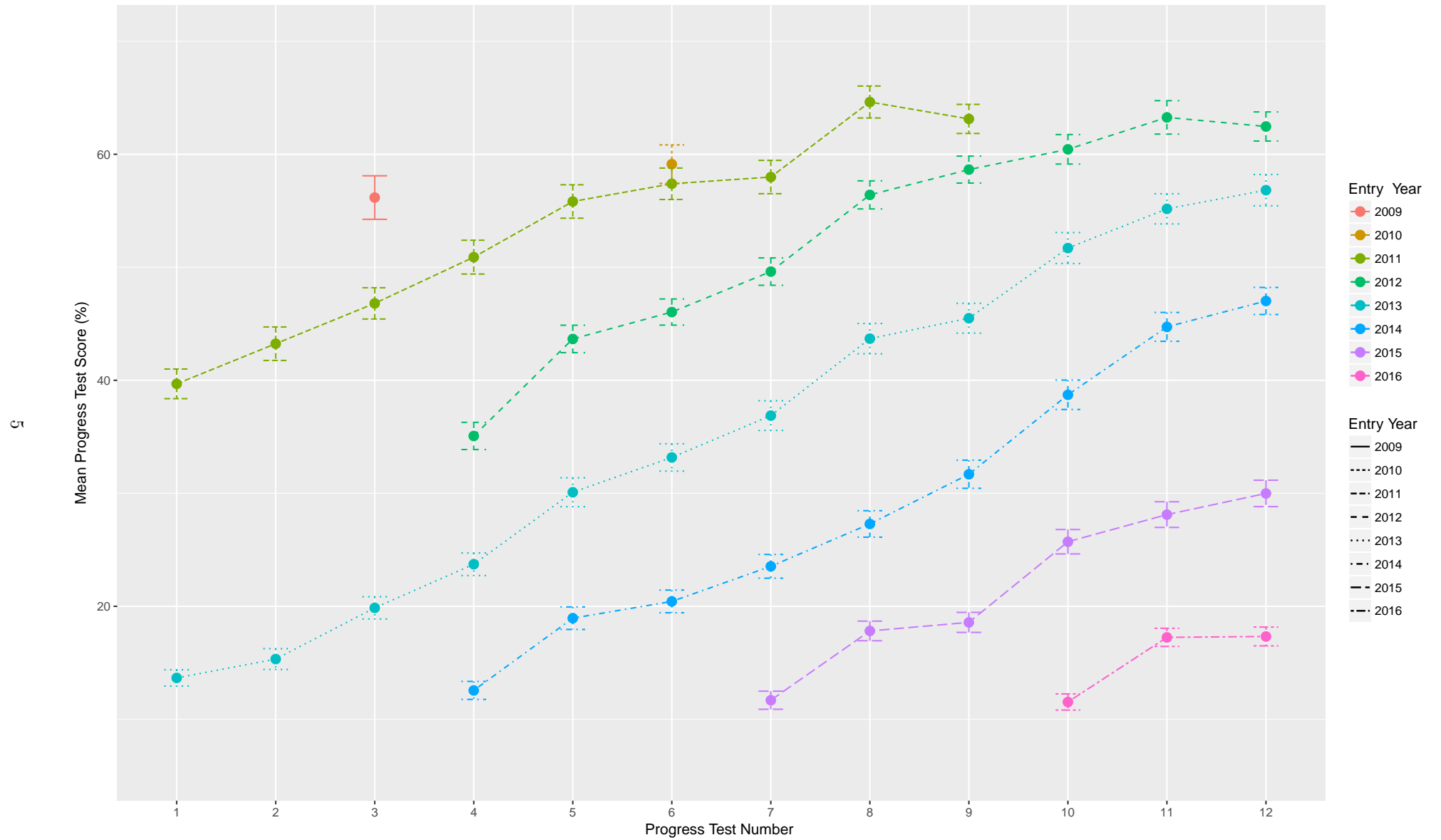
Year 6 students in 2013 and 2014 also sat PT03 ($n = 123$) and PT06 ($n = 138$) as formative tests, respectively. The data from these groups (Cohorts 2009 and 2010) inform the standards that are set for Y6 students, starting with PT07.

The values in Tables 1 and 2 are plotted in Figures 1 and 2.



Date Created: 2016-12-14 09:29:34

Figure 1: Longitudinal Mean Progress Test Scores



Date Created: 2016-12-14 09:29:34

Figure 2: Longitudinal Progress Test Scores (Mean and 95% CI)

3 Test Statistics

3.1 Summary Statistics

Table 3 to Table 7 show the summary statistics of the scores in all progress tests that the current Year 2 to Year 6 students have sat.

Table 3: Year 2 Descriptive Statistics

PT	Minimum	P25	Median	Mean	SD	P75	P95	Maximum	IQR	Range
PT10	-1.8	7.4	10.5	11.5	6.1	14.5	22.3	38.5	7.1	40.3
PT11	2.0	12.1	16.9	17.3	6.8	21.2	29.8	37.7	9.1	35.7
PT12	-0.2	11.9	17.1	17.3	7.0	21.9	30.0	35.8	10.0	36.0

Table 4: Year 3 Descriptive Statistics

PT	Minimum	P25	Median	Mean	SD	P75	P95	Maximum	IQR	Range
PT07	-0.6	6.8	10.8	11.7	6.8	15.3	25.2	43.6	8.5	44.2
PT08	4.5	12.6	17.1	17.8	7.4	22.0	32.5	44.9	9.4	40.4
PT09	-2.0	13.5	17.9	18.6	7.5	23.2	31.5	50.4	9.7	52.4
PT10	3.7	18.6	25.4	25.7	9.2	32.2	41.4	57.0	13.6	53.3
PT11	6.9	21.2	27.6	28.1	9.6	33.7	44.8	57.5	12.5	50.6
PT12	4.6	22.9	28.5	30.0	9.9	35.8	49.2	57.5	12.9	52.9

Table 5: Year 4 Descriptive Statistics

PT	Minimum	P25	Median	Mean	SD	P75	P95	Maximum	IQR	Range
PT04	-2.0	8.5	11.9	12.6	6.5	15.8	23.4	61.9	7.3	63.9
PT05	1.2	12.9	17.6	18.9	8.1	23.7	32.4	65.1	10.8	63.9
PT06	5.6	14.5	19.4	20.4	8.2	25.2	34.4	76.1	10.7	70.5
PT07	7.8	17.6	22.5	23.5	8.5	28.2	38.6	72.2	10.6	64.4
PT08	7.5	20.7	26.2	27.3	9.5	32.6	45.7	75.2	11.9	67.7
PT09	10.9	24.2	30.4	31.7	10.0	38.2	48.1	71.8	14.0	60.9
PT10	15.2	31.6	37.3	38.7	10.3	45.0	59.0	75.4	13.4	60.2
PT11	22.0	37.1	44.3	44.7	10.2	51.0	61.8	77.0	13.9	55.0
PT12	27.3	40.1	46.4	47.0	9.5	52.9	63.4	78.8	12.8	51.5

Table 6: Year 5 Descriptive Statistics

PT	Minimum	P25	Median	Mean	SD	P75	P95	Maximum	IQR	Range
PT01	1.8	9.7	13.0	13.7	5.7	16.7	23.2	36.2	7.0	34.4
PT02	0.4	9.7	14.7	15.3	7.2	19.9	28.3	40.5	10.2	40.1
PT03	3.7	14.4	19.3	19.9	7.7	24.6	33.9	44.3	10.2	40.6
PT04	6.1	18.0	22.7	23.7	7.8	27.7	37.9	53.9	9.7	47.8
PT05	11.8	23.3	29.1	30.1	9.9	34.9	49.6	63.6	11.6	51.8
PT06	12.8	26.4	32.0	33.2	9.4	38.9	50.4	65.6	12.5	52.8
PT07	8.5	29.3	36.2	36.9	10.2	43.3	54.4	69.7	14.0	61.2
PT08	15.2	36.6	43.6	43.7	10.4	49.8	61.5	80.7	13.2	65.5
PT09	23.2	38.5	44.2	45.5	10.2	52.2	65.3	76.0	13.7	52.8
PT10	25.8	44.3	51.3	51.7	10.5	58.6	71.1	84.6	14.3	58.8
PT11	27.6	47.8	53.6	55.2	10.1	61.5	73.8	83.9	13.7	56.3
PT12	29.6	49.5	56.2	56.8	10.4	63.2	76.0	86.5	13.7	56.9

Table 7: Year 6 Descriptive Statistics

PT	Minimum	P25	Median	Mean	SD	P75	P95	Maximum	IQR	Range
PT04	10.9	27.7	34.6	35.1	9.4	42.1	50.6	64.1	14.4	53.2
PT05	18.2	36.2	43.8	43.7	9.4	50.6	59.8	70.5	14.4	52.3
PT06	24.6	40.2	45.9	46.0	8.9	51.3	61.3	78.6	11.1	54.0
PT07	22.2	43.0	49.6	49.6	9.4	55.5	66.1	77.8	12.5	55.6
PT08	33.5	49.2	56.6	56.4	9.6	63.4	72.3	81.7	14.2	48.2
PT09	35.3	51.8	59.1	58.6	9.3	65.0	74.0	82.9	13.2	47.6
PT10	37.9	54.3	59.8	60.4	8.8	67.4	73.9	83.6	13.1	45.7
PT11	29.8	57.3	63.6	63.3	9.6	69.6	79.5	86.9	12.3	57.1
PT12	36.9	55.4	62.5	62.5	9.6	69.3	78.1	86.0	13.9	49.1

The diagram in Figure 3 visually captures five of the summary statistics for PT12 scores that are found in the descriptive statistics tables. The ends of the whiskers correspond to the minimum and maximum PT12 scores. Dots are added when there are outliers. The horizontal line inside the box corresponds to the median PT12 score. The top and bottom borders of the box refer to the 75th percentile (P75) and 25th percentile (P25), respectively. The length of the entire plot (including the outliers) gives us an idea of the range of the scores in the distribution. The length of the box shows us the range of scores of the middle 50% of the distribution, the interquartile range (IQR).

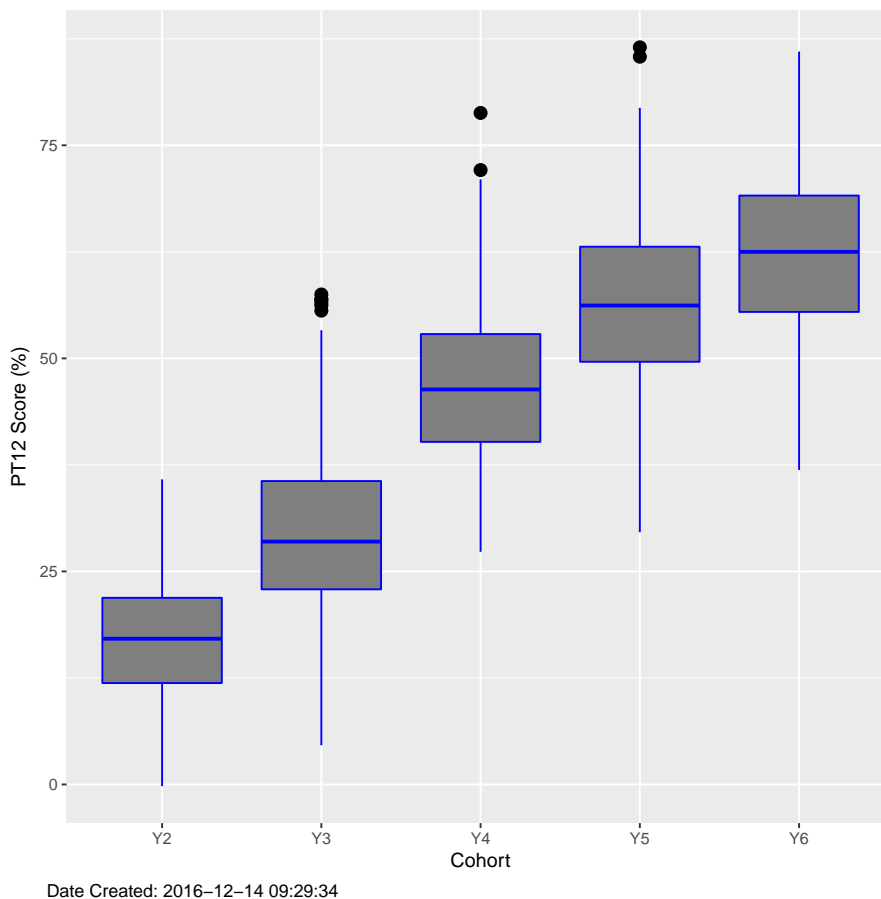


Figure 3: Boxplots of PT12 Scores By Cohort

Table 8: Distribution of students per cohort that will be awarded different PT grades

Cohort	Unsatisfactory	Borderline	Satisfactory	Excellent
Year 2	5%	10%	80%	5%
Year 3	4%	8%	83%	5%
Year 4	3%	6%	86%	5%
Year 5	2%	4%	89%	5%

3.2 Grade Boundaries

Because Year 2 to Year 5 students are graded normatively, certain percentages of the cohort are expected to be classified in one of the four PT grades as shown in Table 8. The cut-scores that would yield PT12 grades with this distribution are found in Table 9.

Table 9: Cut Scores

Cohort	Position	Floor Score	Ceiling Score	Boundary	Cut Score
Y2	13.9	7.1	7.1	UB	7.1
Y2	41.7	9.8	9.8	BS	9.8
Y2	264.1	30.0	30.4	SE	30.2
Y3	11.1	15.0	15.0	UB	15.0
Y3	33.4	19.2	19.2	BS	19.2
Y3	264.1	49.2	49.6	SE	49.4
Y4	7.4	29.4	30.0	UB	29.7
Y4	22.2	35.0	35.4	BS	35.2
Y4	234.6	63.3	63.5	SE	63.4
Y5	4.4	36.0	37.5	UB	36.8
Y5	13.1	41.5	41.5	BS	41.5
Y5	208.0	76.0	76.0	SE	76.0

Year 6 boundary scores were determined using the item-mapping method of standard setting as proposed by Wang (2003). The results were triangulated using historical data as well as Cohen’s method (Cohen-Schotanus and van der Vleuten, 2010). Number-correct scores, instead of formula (or negatively marked) scores were analysed using the Rasch measurement software *Winsteps* (Linacre, 2016) to estimate the locations of the items (*difficulty*) and of test takers (*ability*) on the same scale in the map.

See MacCann and Stanley (2006) for comparison of item mapping with the more familiar methods such as Angoff procedure and the bookmark method.

The two groups of experts that participated in the item-mapping sessions were composed of clinicians familiar with the standard expected of both final year medical students and new graduates. The results of both standard setting sessions were combined to propose that the PT12 score of 41% be the Pass-Fail cut-score.

The item-mapping method is robust for determining the Pass-Fail cut-score. However, experts may find it difficult to find a consensus when determining the Excellent-Pass cut-score and different groups are more likely to propose very different cut-scores. Hence, the Excellent-Pass cut-score, with the approval of the Assessment Subcommittee, was determined using norm-referencing. Just like in other cohorts, the progress test grade that separated the top 5% of the cohort became the cut-score. With the Year 6 cohort, that PT12 score was 78.2%.

3.3 Response Summary

3.3.1 PT12 Data

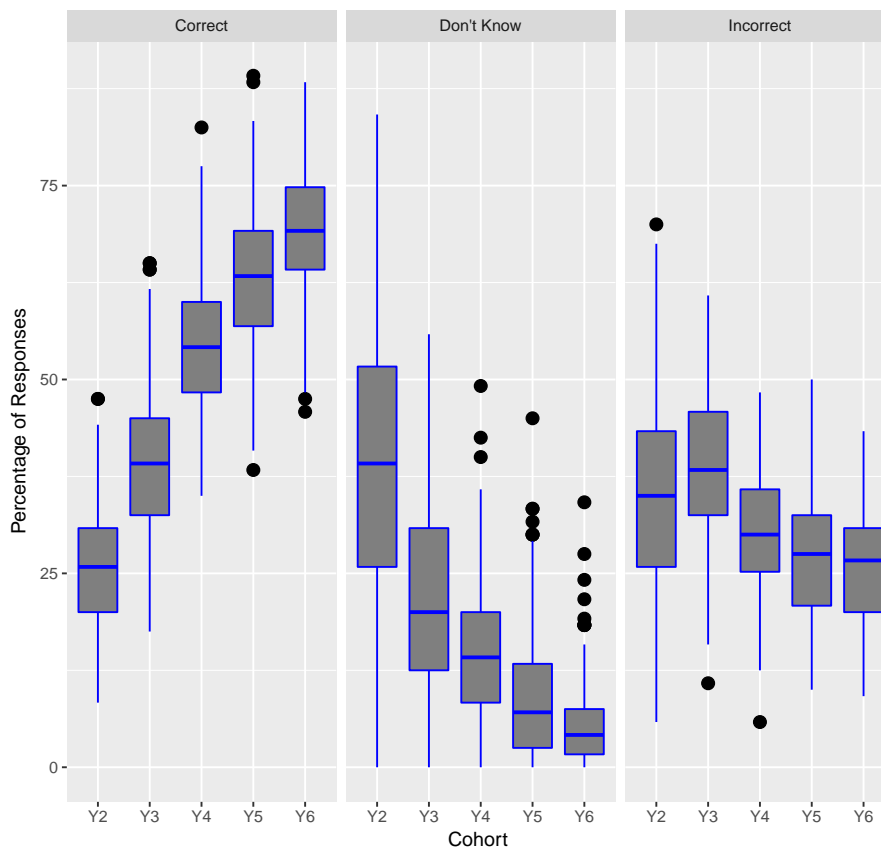
Table 10 shows the average counts of the three possible response types to the 120 PT12 items: correct, incorrect and *don't know* (DK) responses.

In progress testing, students who are relatively new to the programme are, on average, expected to have more DK and fewer correct responses than those who have been in the programme for a longer period of time.

Table 10: Mean Number of Responses to 120 Items

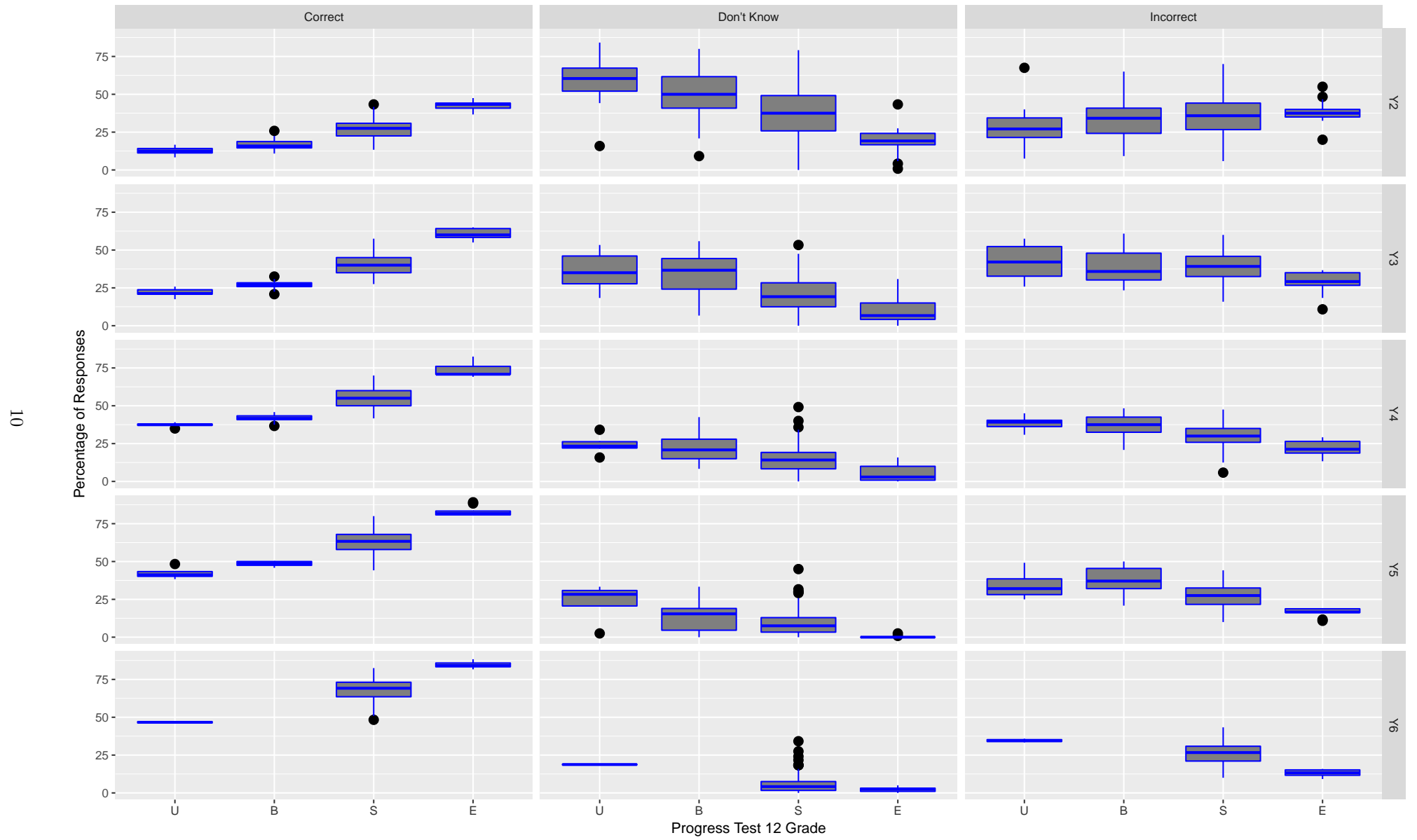
Response	Y2	Y3	Y4	Y5	Y6
Correct	31.5	47.6	65.6	76.4	82.7
Don't Know	45.9	26.0	17.7	11.0	6.4
Incorrect	42.7	46.4	36.7	32.7	30.9

Figure 4 shows the summary statistics *minimum*, *P25*, *median*, *P75* and *maximum* for each response type per cohort. Boxplots that show the same summary statistics for each response type per PT12 grade of each cohort are found in Figure 5.



Date Created: 2016-12-14 09:29:34

Figure 4: Response Summary Boxplots



Date Created: 2016-12-14 09:29:34

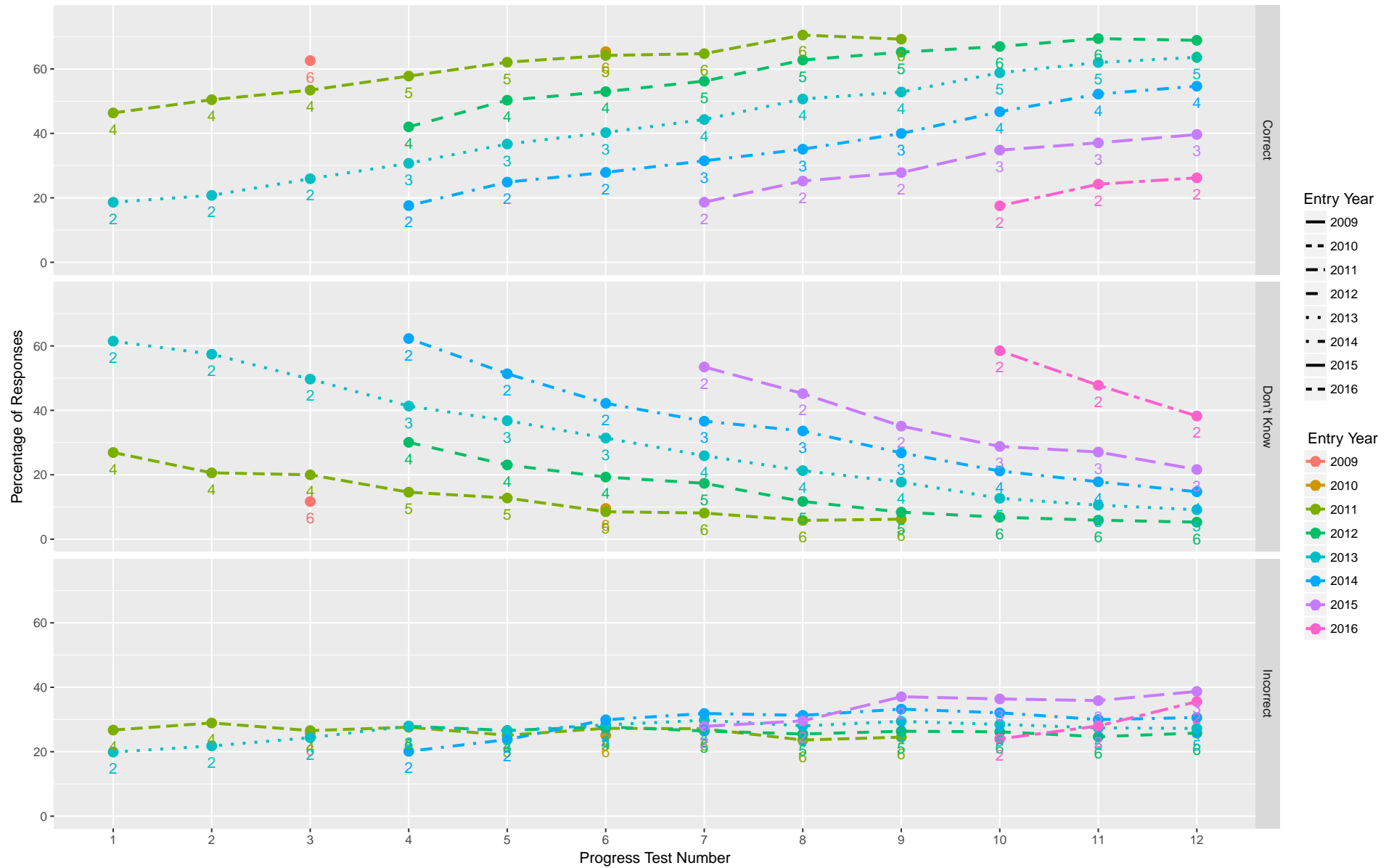
Figure 5: Response Summary Boxplots by Cohort and PT12 Grade

3.3.2 Longitudinal Data

Table 11: Percentage of Responses

Cohort	Response	PT01	PT02	PT03	PT04	PT05	PT06	PT07	PT08	PT09	PT10	PT11	PT12
2016	Correct										18	24	26
2016	Don't Know										58	48	38
2016	Incorrect										24	28	36
2015	Correct							19	25	28	35	37	40
2015	Don't Know							53	45	35	29	27	22
2015	Incorrect							28	30	37	36	36	39
2014	Correct				18	25	28	32	35	40	47	52	55
2014	Don't Know				62	51	42	37	34	27	21	18	15
2014	Incorrect				20	24	30	32	31	33	32	30	31
2013	Correct	19	21	26	31	37	40	44	51	53	59	62	64
2013	Don't Know	61	57	50	41	37	31	26	21	18	13	11	9
2013	Incorrect	20	22	24	28	26	28	30	28	29	28	27	27
2012	Correct				42	50	53	56	63	65	67	69	69
2012	Don't Know				30	23	19	17	12	8	7	6	5
2012	Incorrect				28	27	28	26	25	26	26	25	26
2011	Correct	46	50	53	58	62	64	65	71	69			
2011	Don't Know	27	21	20	15	13	9	8	6	6			
2011	Incorrect	27	29	27	28	25	27	27	24	25			
2010	Correct						65						
2010	Don't Know						10						
2010	Incorrect						25						
2009	Correct			63									
2009	Don't Know			12									
2009	Incorrect			26									

Figure 6 shows the summaries of correct, don't know, and incorrect responses from the first to the current progress test.



Date Created: 2016-12-14 09:29:34

Figure 6: Summaries of Responses

References

- Cohen-Schotanus, J. and van der Vleuten, C. P. M. (2010). A standard setting method with the best performing students as point of reference: Practical and affordable. *Medical Teacher*, 32(2):154–160. <http://dx.doi.org/10.3109/01421590903196979>.
- Linacre, J. (2016). Winsteps (version 3.92.1) [computer software]. www.winsteps.com. Retrieved February 2016.
- MacCann, R. G. and Stanley, G. (2006). The use of Rasch modeling to improve standard setting. *Practical Assessment Research & Evaluation*, 11(2):1–17. Available online: <http://pareonline.net/getvn.asp?v=11&n=2>.
- Wang, N. (2003). Use of the Rasch IRT model in standard setting: An item-mapping method. *Journal of Educational Measurement*, 40(3):231–253. <https://goo.gl/1J0aau>.