

# Checking the STAR norms

CHARLES DARR

In this edition of *Assessment News* we take a close look at the new Supplementary Tests of Achievement in Reading (STAR). In particular, we consider how confident teachers can be in comparing their students' test scores against the national norms that have been published by the New Zealand Council for Educational Research (NZCER).

In recent years the New Zealand Council for Educational Research (NZCER) has been able to work with schools across New Zealand to help them make more use of their assessment data. For instance, the NZCER Marking Service allows schools to access a suite of online reporting and analysis options related to a range of tests. As well as providing increased reporting and analysis opportunities for schools, developing these kinds of services has also allowed NZCER to examine how the tools it produces are performing in real classroom use. Last year, NZCER was able to compare the published norms for the new second edition of the Supplementary Tests of Achievement in Reading with data from schools that have made use of the revised tests. The results show that the published norms accurately reflect norms generated from classroom use.

The new edition of STAR was published at the end of 2011.<sup>1</sup> One feature of the new edition is the ability to report achievement on the same measurement scale regardless of which level of the test has been used to assess the student. This means teachers can select a test at the level of difficulty that best suits their students—regardless of students' year level—and use the scale score to track students' achievement and progress over time. In addition, once students' achievement levels have been located on the scale, they can be compared with the

achievement of nationally representative samples of students at Years 3 to 9. These published norms are based on the scores of students who completed a test as part of the STAR standardisation trial held in March 2011.

For the standardisation trial schools across New Zealand were systematically sampled to ensure that the range of students involved represented the range of achievement present in each year level. In all, over 7,500 students took part in the trial (about a thousand at each year level). The results from the trial were used to calculate norms at each year level. But, how well do these published norms reflect the range of achievement we see when students take STAR tests as part of normal classroom programmes?

NZCER has been able to check the published normative information for STAR by using data processed through the NZCER Marking Service. The Marking Service allows schools to upload their STAR scores to produce a range of reporting. The scores uploaded to the Marking Service in Term 1, 2012 for students in Years 3 to 8 were used to create a set of Marking Service norms for the new STAR. These Marking Service norms could be compared with the published norms generated by the standardisation trial.

Table 1 compares the trial sample with the Marking Service sample. In the table the columns headed "PN" show the number of students who

## ASSESSMENT NEWS

TABLE 1. NUMBERS IN THE PUBLISHED NORMS (PN) AND MARKING SERVICE (MS) NORM GROUPS BY QUINTILE

Quintile	Student year level													
	Year 3		Year 4		Year 5		Year 6		Year 7		Year 8		Total	
	PN	MS	PN	MS	PN	MS	PN	MS	PN	MS	PN	MS	PN	MS
1	166	1280	174	1245	210	1208	233	1202	267	738	288	734	1338	6407
2	253	1279	256	1161	246	1042	273	1189	196	1960	235	1765	1459	8396
3	279	1622	322	1569	195	1559	175	1519	255	1961	229	1698	1455	9928
4	168	1997	206	1663	231	1515	253	1375	119	2094	103	1459	1080	10103
5	201	3075	164	2563	290	2235	294	2230	150	1504	113	1391	1212	12998
Total	1067	9253	1122	8201	1172	7559	1228	7515	987	8257	968	7047	6544	47832

made up the sample used to construct the published norms (PN). The columns headed “MS” shows the number of student scores available from the Marking Service (MS). The numbers in the table are broken down by quintile. A quintile is a pairing of school deciles. Quintile 1 includes decile 1 and 2 schools, quintile 2 includes decile 3 and 4 schools and so on. When preparing norms it is important to make sure each quintile is properly represented.

As can be seen, the Marking Service data set is much larger than the original trial data set. Each quintile is well represented, however the proportion of higher quintile schools is greater in the Marking Service sample than in the original trial sample.

Figure 1 uses box plots to compare the published norms with the achievement of students in the Marking Service sample. At each year level the box plots are presented in pairs, with the first plot in each pair used to represent the achievement of students in the Marking Service sample. At each year level the two box plots describe very similar distributions of achievement. It is likely that any differences can be explained by the fact that students from higher quintile schools are slightly over-represented in the Marking Service data. This would tend to lift their achievement compared to a more representative norm group.

The distribution of achievement seems to differ slightly more at Year 3 than at other year levels. The difference is not great and there is not yet sufficient evidence to suggest that the slightly less representative Marking Service data can provide a more accurate measure. However, if the achievement of Year 3 students in 2013 is seen to follow the same pattern, NZCER will consider an update of the Year 3 published norm.

Overall, the exercise has shown that the published norms for STAR strongly agree with norms generated from the data we have from schools that are making use

of the tests as part of their normal classroom practice. When schools want to use the revised STAR to make normative comparisons they can be confident that the published norms for the new tests accurately reflect achievement norms for New Zealand students.

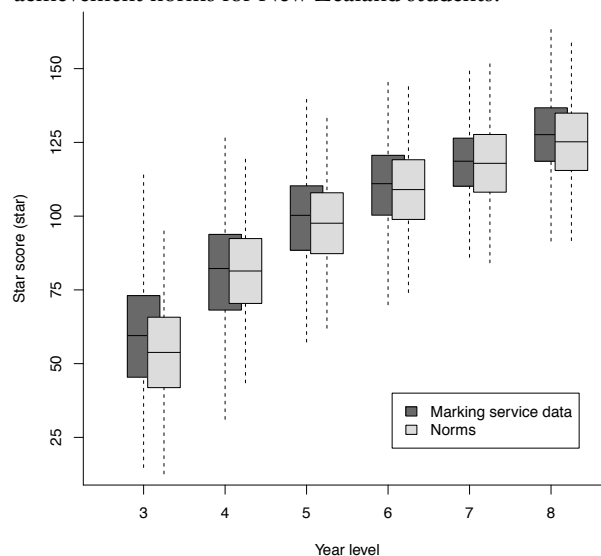


FIGURE 1. COMPARISON OF NORMS BASED ON MARKING SERVICE DATA WITH PUBLISHED NORMS FOR STAR

### Note

- 1 STAR was originally developed in the early 2000s for NZCER by Dr Warwick Elley. Dr Elley also led the development of the revision in collaboration with researchers from NZCER.

▶ **CHARLES DARR** is manager of assessment design and reporting at the New Zealand Council for Educational Research.

Email: [charles.darr@nzcer.org.nz](mailto:charles.darr@nzcer.org.nz)