Introduction

A study was carried out that compared Programme for International Student Assessment (PISA) 2003 mathematics results with Trends in International Mathematics and Science Study (TIMSS) 2003 grade 8 Mathematics results, using country mean scores for 22 participating countries of both studies (Wu, 2009). It was found that Western countries generally performed better in PISA than in TIMSS, and Eastern European and Asian countries generally performed better in TIMSS than in PISA. Furthermore, two factors, content balance and years of schooling, accounted for most of the variation in PISA country mean scores after controlling for TIMSS country mean score. Consequently, the rankings of countries in the two studies can be reconciled to a reasonable degree of accuracy. Further, the Organisation for Economic Co-operation and Development (OECD) publication, *Learning mathematics for life* (OECD, 2009, chap. 4), showed that the relative performances of countries by traditional mathematics content areas (Algebra, Data, Geometry, Measurement and Number) differ significantly, and that countries are grouped according to their similarities in students' responses to the items. Gronmo and Olsen (2008) also showed that the compositions of the PISA and TIMSS tests differ considerably in terms of mathematics content and item format. They concluded that content differences between PISA and TIMSS contributed to the observed differences in the rankings of countries in PISA and TIMSS.

The fact that content balance has a significant effect on country performance suggests that students in different countries have particular relative strengths and weaknesses. If these specific strengths and weaknesses are identified beyond the level of broad content categories, mathematics educators in each country can be informed of the specific skills students have or lack. This will also provide a further insight into what PISA and TIMSS are each assessing, beyond the usual rhetoric about curriculum-based and non-curriculum-based focus.

This paper attempts to examine item level skills in order to identify strengths and weaknesses, thus providing specific instructional feedback to mathematics educators as well as test designers. The item analysis is based on data from six countries: Australia, England, United States, Hong Kong, Taiwan and Korea. The results for the three Western countries show a similar pattern, as do the results for the three Asian countries. There is strong evidence that cultural traditions of learning mathematics have an impact on students' performances on different types of test items.

PISA and TIMSS Mathematics Frameworks and Tests

Comparisons of PISA and TIMSS mathematics frameworks can be found in a number of publications (American Institutes for Research, 2005; Hutchison & Schagen, 2007; National Center for Education Statistics, 2008; Neidorf, Binkley, Gattis, & Nohara, 2006). These comparisons tend to produce a descriptive list of similarities and differences between the two published frameworks, such as the classifications of the content domains and the cognitive domains. However, few published comparisons critically examined the differences. For example, would one framework lead to a test that is completely different from a test based on the other framework? Is one framework a subset of the other? If so, what is missing in that framework? Are the two frameworks essentially the same, other than nomenclature of the classifications? This paper looks at these issues and, in doing so, it is hoped that a better understanding is gained in relation to key differences between PISA and TIMSS at the level of the tests and items, not just at the level of broad aims and orientations, as reflected in the frameworks.

At a first glance, most would conclude that both PISA and TIMSS mathematics frameworks are comprehensive. It does not appear that one framework is necessarily a subset of the other, or that something is glaringly missing from either framework. However, the PISA mathematics framework suggests its more inclusive approach, with the following line at the beginning of the framework document:

> *Rather than being limited to the curriculum content students have learned, the assessments focus on determining if students can use what they have learned in the situations they are likely to encounter in their daily lives. (OECD, 2003, p. 24)*

The word “limited” suggests that PISA is attempting to be more inclusive in terms of coverage of the mathematics domain. It also suggests that the school curriculum, whether intended, implemented, or attained, does not focus on whether students can use what they have learned. Does this mean that TIMSS, being more curriculum based, does not assess whether students can use what they have learned? The TIMSS mathematics framework states the following:

> *While the assessment of abilities such as solving nonroutine problems and reasoning with mathematics will be of special interest, the factual, procedural, and conceptual knowledge*