Translating test items into Norwegian – without getting lost in translation?

REIDAR MOSVOLD, JANNE FAUSKANGER,
ARNE JAKOBSEN AND KJERSTI MELHUS

In relation to the Learning Mathematics for Teaching (LMT) project, sets of measures were created in order to analyse teachers Mathematical Knowledge for Teaching (MKT). This article presents some of the challenges and complexities involved in an attempt to translate and adapt these measures for use with Norwegian teachers. The measures were originally created for use in a U.S. context only, and a number of differences between the two countries contribute to increase the difficulty of doing this. Our study builds upon a similar Irish study, and this article points to some similar and several additional issues that arise when attempting to translate and adapt the measures for use in Norway.

In mathematics education, there is a growing interest concerning the kinds of knowledge that teachers need in order to teach mathematics or to become effective mathematics teachers (Kotsopoulos & Lavigne, 2008; Davis & Simmt, 2006). For in-service education, the question: "What knowledge do teachers need to become effective teachers?" is important in the process of developing future in-service training.

The *Norwegian ministry of education and research* (KD, 2008a) underlines that teachers' knowledge is important. Still, Norwegian mathematics teachers have less study points (ECTS) than the international average, and they participate in relevant in-service education to a strikingly small extent (UFD, 2005; Grønmo et al., 2004). Research from the last 15 years shows that (U.S.) teachers do not know enough mathematics (Ma, 1999), and as a consequence the students do not learn enough (Ball, Hill &

Reidar Mosvold, *University of Stavanger*
Janne Fauskanger, *University of Stavanger*
Arne Jakobsen, *University of Stavanger*
Kjersti Melhus, *University of Stavanger*

Mosvold, R., Fauskanger, J., Jakobsen, A. & Melhus, K. (2009). Translating test items into Norwegian – without getting lost in translation? *Nordic Studies in Mathematics Education*, 14 (4), 9–31.

Bass, 2005). When analysing 700 teachers in 1st and 3rd grade, researchers found that teachers' knowledge have an effect on the knowledge of their students (Hill, Rowan & Ball, 2005). Falch and Naper (2008) present similar results from Norwegian teachers and their students in lower secondary school. Although research indicates that teachers' knowledge has a positive influence on students' learning, and the slogan: "teachers matter" (UFD, 2002; OECD, 2005) is widely used, it is far from obvious what the content of this knowledge is. There are also no clear guidelines regarding the intended focus for in-service education, at least in Norway (KD, 2008b).

The importance of future research focusing on teachers' knowledge and implications for in-service education is underlined in a report from the U.S. National Mathematics Advisory Panel (Faulkner et al., 2008). Over the years there have been several attempts of investigating the content of teachers' knowledge, and researchers have approached this in different ways (Hill, Sleep, Lewis & Ball, 2007). In the Learning mathematics for teaching (LMT) project, researchers at the University of Michigan developed measures in order to investigate the kind of mathematical knowledge that is needed for teaching (referred to as MKT). They describe this as the kind of mathematical knowledge that is used by teachers. This knowledge is used in the classroom setting, and it is normally related to an overall aim of increasing student's performance in mathematics (Hill, Ball & Shilling, 2008) or as the kind of mathematical knowledge that teachers need to carry out the work of teaching mathematics (Ball, Thames & Phelps, 2008). These measures were developed for use in the U.S., and unlike international student assessments like TIMSS (e.g. Mullis, Martin, Gonzales & Chrostowski, 2004) and PISA (e.g. OECD, 2004), these measures were never intended to be used outside of the U.S. As a result of this, many of the items contain contexts that might be specific to the U.S., and the mathematical content was not made to fit with curricula in other countries. A process of translating and adapting these measures for use in another country therefore involves several problematic issues. Still, a large amount of money have been used to develop the items in the U.S., and we found it interesting to investigate whether a translation and adaptation of the items into Norwegian would be fruitful or even possible. Another reason for going into such a project is that MKT items have been used in studies like TEDS-M, but little or no efforts appear to have been made in order to discuss or analyse possible issues related to a translation and adaptation of such items. Questions regarding translation are often answered with reference to the fact that professional translators have been used. We believe that it is important for us as researchers within the field of mathematics education to analyse

and discuss problematic issues related to translation when using items and tests like these, and we therefore decided that it was vital for us to go into the translation ourselves rather than leave this to professional translators. Besides, a translation of the MKT items is not (only) about making a good translation of the text itself, but it is very much a matter of representing the mathematical and pedagogical contents of the items in a correct and meaningful way. By leaving professional translators with the full responsibility for this process, we open up to possible problems that might become serious threats to the validity of the study.

This article represents an attempt to identify and discuss issues that arise when translating MKT measures. Our research question is:

> What problems occur in the process of translating and adapting the MKT measures from a U.S. context into a Norwegian context?

In answering this question, we are building upon a similar study that was carried out in Ireland (cf. Delaney et al., 2008). The Irish study provides a set of steps that are recommended for researchers who attempt to adapt the measures from one country to another.

Theoretical foundations

This study follows the tradition that has evolved as an expansion of Shulman's (1986) concept of pedagogical content knowledge to the more specialised knowledge that is required for teachers of mathematics. In addition to describe the theoretical background of this tradition, we have to pay attention to the theoretical assumptions and issues that are involved in the process of translating and adapting the measures.

Mathematical knowledge for teaching

Some years ago there was a widespread opinion that if teachers knew enough mathematics, their teaching would be good and their students would learn mathematics. The content of in-service education then became purely mathematical (Cooney, 1999). On the other extreme, there appeared to be a consensus in some Norwegian teacher colleges that it was possible to become an effective mathematics teacher without knowing much mathematics (Haaland & Reikerås, 2005). Begle (1968) and Eisenberg (1977) argued that effective teaching is about more than the teachers' mathematical competence. Shulman (1986) addressed four questions, one of which is "what are the sources of the knowledge base for teaching?" He tried to put teacher knowledge into certain categories: subject matter knowledge or content knowledge, pedagogical content

knowledge, and knowledge of curriculum. These headings pointed to the fact that mathematical knowledge alone does not automatically transfer into more effective teaching.

Researchers in the LMT project based their work on Shulman's, and they tried to identify and specify the mathematical knowledge that teachers need. This knowledge not only includes aspects of pedagogical content knowledge, but also incorporates subject matter knowledge, both common and specialised to the work of teaching. The researchers seek to understand and measure MKT. A proposed model of the construct of MKT can be seen in figure 1.

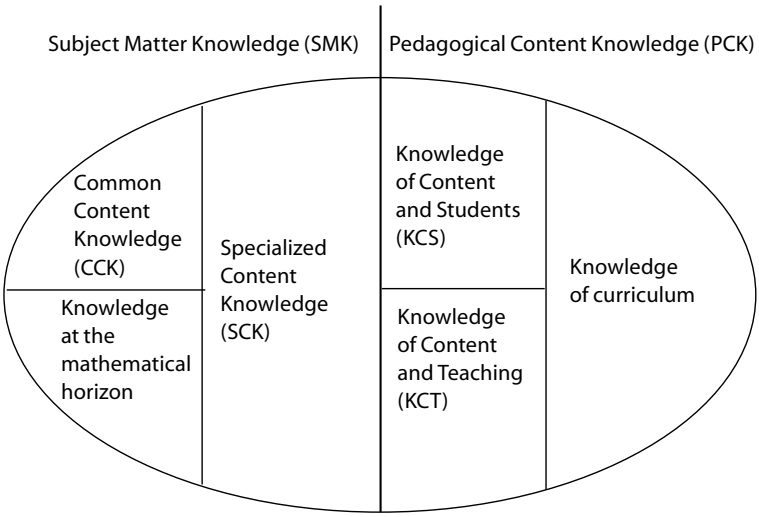

Figure 1. *Mathematical knowledge for teaching (based on Hill, Ball & Schilling, 2008, p. 377).*

The figure shows the correspondence between the researchers' current map of the domain MKT and Shulman's (1986) current categories: subject matter knowledge and pedagogical content knowledge. Shulman's third category, knowledge of curriculum, is placed within pedagogical content knowledge.

The left side of the oval contains two strands that lie outside Shulman's pedagogical content knowledge: common content knowledge and specialised content knowledge. Common content knowledge is knowledge that is used in the work of teaching, in ways that correspond with how it is used in other professions or occupations that also makes use of mathematics. Specialised content knowledge is the mathematical knowledge "that allows teachers to engage in particular teaching tasks, including

how to accurately represent mathematical ideas, provide mathematical explanations for common rules and procedures, and examine and understand unusual methods to problems" (Hill, Ball & Shilling, 2008, p. 378). Common content knowledge is similar to Shulman's subject matter knowledge, whereas specialised content knowledge is a more recent conceptualisation. Both are mathematical knowledge. The right side of the oval contains knowledge of content and students, knowledge of content and teaching and knowledge of curriculum. All three were included in what Shulman referred to as pedagogical content knowledge, and this again is a subset of the larger construct: MKT. Horizon knowledge is "an awareness of how mathematical topics are related over the span of mathematics included in the curriculum" (Ball, Thames & Phelps, 2008, p. 42).

In a recent article, Hill, Ball and Schilling (2008) describe an effort to conceptualise and develop measures of teachers' combined knowledge of content and students. The authors point to a widespread agreement that effective teachers have a unique knowledge of students' mathematical ideas and thinking, but too few studies have focused on conceptualising this domain and on measuring this knowledge. Although the domains presented in figure 1 have been identified in the U.S., the domains of knowledge may differ in other settings, such as Norwegian settings. Measuring teachers' MKT is not straightforward. Since the researchers have put a lot of time, money and effort into the development of the MKT measures, it would be interesting to translate, adapt and use them in other countries. However, such a translation of measures might bring a lot of aspects into question. One example relates to the meaning of mathematical practice. Terms can be misunderstood within the same language and cultural context, and when we translate the term into a different language there might be even more room for misunderstanding (Delaney et al., 2008).

Even if we are building our research on the theoretical framework of MKT, it is important to note that other researchers' perspective entails different and complementary foci (e.g. Schoenfeld, 2007; Silverman & Thompson, 2008; Thompson, Carlson & Silverman, 2007). These critics will not be further discussed in this article.

Lost in translation

According to the PISA 2003 Technical Report (Adams, 2005), translation errors are known to be a major reason why some items function poorly in international tests. Regarding MKT, studies provide little information as to how measurement instruments are adapted for use outside the U.S., and in the different publications little information is given about translation issues arising in the research (Delaney et al., 2008). Ma (1999) has

for example compared U.S. and Chinese teachers' knowledge of mathematics, but she provide little information as to how the measurement instruments were adapted, and little information is given about translation issues that arose in the research. This is critical "because misunderstandings of terms can alter whether and how instruments discern teacher knowledge" (Delaney et al., 2008, p.5). Even before translating the MKT measures we have to be aware of the fact that multiple-choice measures are not widely used in Norway. This may cause validity problems. It is conceivable that in a culture where multiple-choice formats are unfamiliar, one may have to change the format. But changing the format may be problematic, because it could influence the item's level of difficulty (ibid.). It could also make the item more or less discriminating or change how effectively the item measures the underlying construct. We have decided to keep the multiple choice format for now and evaluate the matter after the pilot study.

Translating the MKT measures into Norwegian is not only a matter of translation from one language to another. It is also a matter of translating and adapting a set of measures that was originally developed for use in one cultural context into a different context. According to Peña (2007), methodological norms cannot easily be translated. If these norms were developed for use in a particular country, they should not only be translated but also adapted to the country or population in target. MKT is a practice-based construct, as it is grounded in the practice of teaching. The basis of the U.S. construct of MKT is the knowledge that is demanded in a U.S. teaching practice. The construct has been developed by systematically studying records of mathematics teaching (e.g. videotapes of lessons, copies of student work, teacher's plans and reflections) to identify the mathematical demands of teaching (Ball & Bass, 2003). Adapting measures developed in the U.S. for use in Norway is not a trivial matter. According to Delaney and colleagues (2008), the MKT items are different from other kinds of item translation. One difference is that the items were not initially designed to be used in other settings. Stiegler and Hiebert (1999) suggest that the work of teaching is different in different countries, and if the work of teaching in the U.S. is different from that in Norway, an instrument to measure knowledge for teaching needs to be sensitive to such differences. A second reason is that the MKT items are not grounded in the discipline of mathematics, but in the practice of teaching mathematics. This points out a need to recruit experts in the practice of teaching in Norway in the process of adapting the items. The third point is the actual and potential areas of difference in MKT across countries related to teachers, students, mathematics and teaching materials (Delaney et al., 2008).

Geisinger (1994) suggests that tests should continue to measure the same characteristics as was intended. The content of the test should also remain the same. An important methodological goal for translating the MKT measures into Norwegian therefore is to ensure equivalence at the level of context and opportunity. An attempt to adapt the U.S. measures to an Irish context (cf. Delaney, 2008) emphasised the need to establish whether the MKT construct is equivalent in different settings. Construct equivalence is thus an important aspect of the validation process.

Various terms are used in cross-cultural research to describe different aspects of equivalence. According to Johnson (1998), the terms are not always well-defined. There might also be serious overlap between these terms. The threats against validity are serious, and Peña (2007) claims that it is not sufficient to use certain translation techniques in order to establish linguistic equivalence. Attaining a high quality in the translation of the MKT measures is therefore not enough, and it does not ensure equal opportunities for Norwegian teachers to demonstrate their MKT. The type of equivalence identified as necessary depends on the goals of the study. If issues related to measurement are overlooked, inferential errors might occur, and Singh (1995) underlines that few empirical studies take this seriously. The effects are not only complex and unpredictable, but they might have an influence on everything. Following the steps described by Singh, Delaney (2008) studied three aspects of construct equivalence: functional equivalence, conceptual equivalence and instrument equivalence, before using the MKT measures to collect data to learn more about the Irish teachers MKT. We build our research on his work, but Delaney only focused on the translation from American English into British English, so when we have to make a translation into a completely different language, we also have to focus on what Peña (2007) calls linguistic equivalence (see figure 2).

Functional equivalence relates to whether or not the MKT construct serves the same function in all countries. In order for students to acquire knowledge, the teacher must have some kind of knowledge related to teaching (in this case, MKT). This construct – MKT – has a universal function, and thus satisfies the requirements of having functional equivalence (cf. Delaney, 2008).

Two important questions related to conceptual equivalence are:

- Does the construct of MKT mean the same in Norway as in the U.S.?
- Are the demands for primary school mathematics teaching in Norway similar to the knowledge conceptualised in the U.S. construct of MKT?

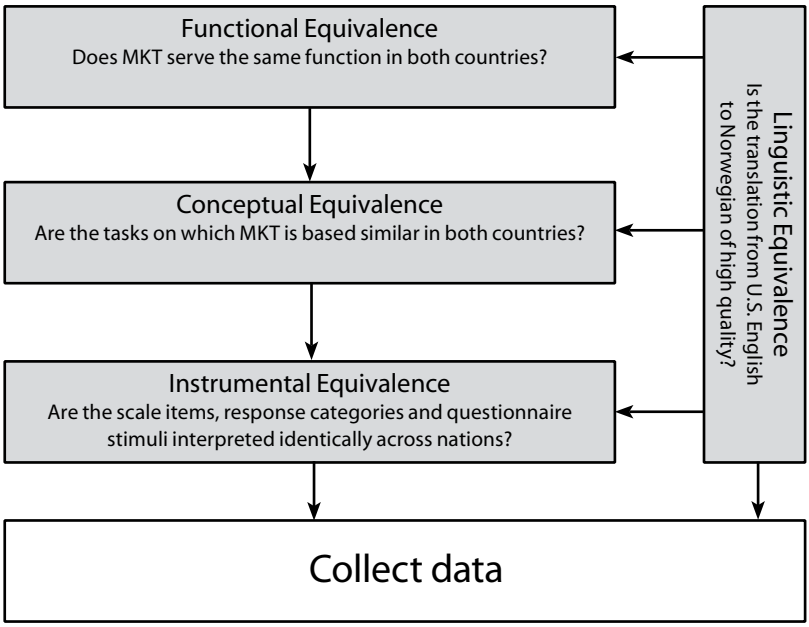

Figure 2. *Steps in establishing construct equivalence. (Further development of a figure from Singh (1995))*

To answer similar questions in an Irish context, Delaney (2008) examined the construct more closely by studying the work of teaching in Ireland. He compared that work to conceptions of the work of teaching that informed the development of MKT. Delaney also studied literature about the construct, and he analysed items based on the construct. He found relatively minor differences in this analysis. One possible explanation might be that these two countries share a common language. This could make it easier for ideas and conceptions about teaching to travel back and forth between Ireland and the U.S.

Norway and the U.S. do not share a common language, so it is possible that more differences may emerge if the tasks that informed the MKT were compared to tasks of teaching in Norway. Since we have to take into account the added complexity of a different language, attempts to ensure conceptual equivalence will be important in our work. If we cannot assure that concepts are understood in the same way in Norway as in the U.S., the results from our study would be difficult or even impossible to interpret.

Instrument equivalence is related to both the format and the contents of the items. If the multiple-choice items are equally interpreted in

Norway and the U.S., we have instrument equivalence (cf. Delaney, 2008, referring to Singh, 1995). In an early phase of our attempt to adapt the items for use in a Norwegian study, we used a more qualitative approach to describe and document the changes that were made. Some Norwegian teachers agreed to complete the survey under our supervision, and this session was followed up with focus group interviews. One of the challenges that has been given to the teachers in these focus group interviews was to propose changes where necessary, in order to make the items sound realistic to Norwegian teachers. Through this process we hope to find out if construct equivalence exists in how the items are interpreted by Norwegian and U.S. teachers. In the following, however, we focus mainly on the aspect of instrument equivalence related to the translation from American English into Norwegian.

Documenting changes – results

The MKT measures consist of several multiple-choice items, and each complete survey consists of items that cover different areas of school mathematics. We decided to start with a focus on the items that were created for use with teachers in elementary school, and we chose the surveys from 2004, which were the most recent in that category. In 2004, two complete surveys were developed. Each survey consisted of one set of items related to numbers and operations, geometry, and algebra and patterns. We wanted to see how the entire tests worked out in a Norwegian setting, and we therefore decided to translate, adapt and try out a complete test instead of making a selection of items, like they did in Ireland (cf. Delaney et al., 2008).

When translating test items, it is important to ensure linguistic equivalence. Since single translation of test items has proved to be the least trustworthy method, we have used the recommended double translation procedure (Adams, 2005). Double translation means two independent translations from the source language, with reconciliation by a third person. The PISA items go through a process of double translation from two source languages (English and French). According to the PISA 2003 Technical Report (ibid), double translation from the English source only appeared to be effective when accompanied by extensive cross-examination against the French source. In our double translation of MKT items, two pairs translated each item independently, and then compared and discussed the two translations. We had only an English source available, and are aware that double translation from only one source language may be less effective.

In translating the PISA items the main criteria used to recruit leading translators were: Native command of the target language; Professional experience as translators; Sufficient command of the second source language; Familiarity with mathematics; Experience as teachers and/or higher education degrees (Adams, 2005). Our research group satisfy most of these criteria, except professional experience as translators.

Throughout the translation process, we carefully documented all changes that were made to the items (other than direct translation from U.S. English into Norwegian). This was done because we suspect that these changes might influence the teachers' responses to the items. Delaney and colleagues (2008) summarised their changes in the following categories:

1. Changes related to the general cultural context.
2. Changes related to the school cultural context.
3. Changes related to mathematical substance.
4. Other changes.

We decided to use these categories in our own translation process. This was partly because Delaney and colleagues recommended their own results as working guidelines for others who attempt to adapt the items. They included altering spellings to reflect differences between American and British English in category 1 above (changes related to the general cultural context), but we decided to have the translation from U.S. English into Norwegian as a separate category.

The translation from American English into Norwegian was far more complex than the process of translating the items into British English. We therefore had to develop the list of categories further, and we ended up with two new, in addition to the original four categories above:

1. Changes related to the translation from American English into Norwegian in this particular context.
2. Changes related to political directives.

The first of these additional categories have replaced the original sub-category of changes related to spelling in Delaney's first category. The second additional category has been added as a new main category, because we believe this represents some important types of changes that are different from the original categories. The first of these two additional categories were rather obvious, whereas the second emerged somewhat later in the translation process. The category is related to some directives from

the *Norwegian ministry of education and research*, where they have decided to replace certain common terms with new ones. One term, which is frequently used in the original set of items, is class. When talking about schools, it is hard to avoid the term class. The problem is that the *Norwegian ministry of education and research* has decided that we should no longer refer to class in Norway, but rather group of pupils. This is related to a wish to change the way teachers organise their pupils in relation to learning activities. Now, the problematic issue here is that the word *group* is also used in relation to group-work as a method, and there is a potential danger of mixing the two and thereby making some items more complicated to grasp.

This decision to use the word *group* instead of *class* in Norwegian schools might appear trivial, but there is more to it than what we have described above. It appears that a large number of teachers continue to use the word *class*, although the Ministry has decided to avoid it, and for these teachers the word *group* as a replacement to *class* could be both confusing and misleading. If we decided to go for the traditional term *class*, which is no longer the officially correct term, we would probably be faced with a large number of teachers who would argue that our measures were not up to date, and not in line with the official guidelines.

Taking some examples from the items (and we have to use examples from the released items rather than the actual items!), we are going to illustrate some of the types of changes that were made and our concerns and discussions in relation to them.

Below is an example[1] of one of the items that involved several types of changes:

> 10. Students in Mr. Hayes' class have been working on putting decimals in order. Three students – Andy, Clara, and Keisha – presented 1.1, 12, 48, 102, 31.3, .676 as decimals ordered from least to greatest. What error are these students making? (Mark ONE answer.)
>
> a) They are ignoring place value.
> b) They are ignoring the decimal point.
> c) They are guessing.
> d) They have forgotten their numbers between 0 and 1.
> e) They are making all of the above errors.

This is the same item in our translation:

> 10. Elevene til Hans har arbeidet med å sortere desimaltall i stigende rekkefølge. Tre av elevene, Anders, Klara og Kristin, sorterte desimaltall slik:
> 1,1 12 48 102 31,3 0,676

Hvilken feil er det disse elevene gjør? (Marker ETT svar.)

a) De ignorerer plassverdi/posisjonsverdi.
b) De ignorerer desimalkomma.
c) De gjetter.
d) De har glemt at det fins tall mellom 0 og 1.
e) De gjør alle feilene ovenfor.

As we worked on the translation of the items, colour codes were used to identify which words or phrases were changed according to the different categories. These codes are difficult to show in black and white, so they are omitted here.

We discussed several issues in relation to the translation of this item. First, Norwegian students are referred to as *pupils* (or *elever* in Norwegian) as long as they are in compulsory school, and *students* when they enter university. A second issue is that the original sentence referred to the students in Mr. Hayes' *class*, and as described above, you are not supposed to refer to a class of students in Norway. We also changed the name from Mr. Hayes to Hans, which is a common Norwegian first name (further discussions regarding change of names are made in relation to the discussion of general contextual changes below). The passage about putting decimals in order was also discussed. It is more common to sort (*sortere* in Norwegian) numbers than to put them in order, and we talk about decimal numbers (*desimaltall*) rather than decimals. To make this passage sound better in Norwegian, we moved some of the information from the second sentence to the first. As a result, it seems as if Mr. Hayes' students only worked with ordering decimals from least to greatest, whereas the original idea might have been that they worked with ordering decimals in different ways. The first sentences in the stem of this item were difficult to translate directly into Norwegian, and we decided to rewrite them somewhat. When doing this, there is always a possibility of interpreting the sentences in a way that has removed or added information to the item.

In Norway, we use a decimal comma rather than a decimal point, and since comma was used to separate the different numbers that were presented, we had to change this to avoid confusion. One possibility could be to represent the numbers like this: 1,1 – 12 – 48 – 102 – 31,3 – 0,676. From a linguistic point of view, this might be a proper solution, but in a mathematical setting there might be a danger of confusing the – with a subtraction sign. We also discussed the possibility of using semi-colon instead of comma to distinguish the numbers, but we decided that this would result in too much clutter. We therefore ended up presenting the numbers on a separate line with extra space between them. In addition,

we changed .676 into 0,676 because decimal numbers lower than one are never written without the zero in Norwegian. In retrospect we acknowledge that this may not have been the best of choices. By ignoring the comma you get 0676 in the Norwegian version and not 676 as in the US version, making it more unlikely to reach the correct answer b).

In the alternative solutions, we spent some time discussing alternatives a) and d). In a), there is a reference to place value, and we might use the similar word *plassverdi* in Norwegian. Several teachers would rather prefer to use *posisjonsverdi* instead, and we decided to include both alternatives to avoid confusion. Alternative d) was even more problematic to translate. When Americans talk about forgetting your (or their) numbers, this is hard to translate directly into Norwegian. Our translation therefore had to be an interpretation rather than a direct translation. After some discussion, we agreed that the meaning of this sentence must be that the pupils have forgotten that there are numbers between 0 and 1. Another interpretation might be that they did not know this, and a translation into Norwegian might then be: *De kan ikke tallene mellom 0 og 1* (They don't know the numbers between 0 and 1). Such a translation might, however, indicate that the pupils have never been taught this, and we believe that this is not the correct understanding of this alternative solution.

In the following, we use Delaney's categories. The first category concerns changes related to the general cultural context. Examples are given in table 1. In the Irish translation this included changing people's names, making changes related to non-mathematical language and to activities. Delaney and colleagues (2008) included a type of change concerned spelling in this category, because there are some differences of spelling that are specific to American English compared with British English. When making a translation to a different language, like Norwegian, this subcategory is not relevant. As mentioned above, we have added another category related to changes regarding the translation from English into Norwegian in this particular context, which is more relevant here. This category has been placed as a sub-category in their fourth category concerning other changes that were made.

When adapting the MKT measures to an Irish context, changes were made to make the names sound more familiar to Irish teachers. Mr. Ives could therefore be changed into Mr. Fennely, which is a more common Irish name. In Norway, it is common to address teachers with their first name only. This might vary somewhat according to the teachers' age and the level in which they teach, but in primary and lower secondary school the pupils would normally address their teacher as John rather than Mr. (John) Wilson. Mr. Stone would therefore not be addressed by his pupils

Table 1. *Examples of general contextual changes made to items*

Type of change	Example from original U. S. form	Example from our translation
People's names	Mr. Stone	Steinar
	Ms. Wilson	Marianne
Non-mathematical language	batch	porsjon
	cookie	kjeks
	fund-raiser	Redd Barna
	school candy sale	butikken
	M&Ms	seigmenn
Activities	bake cookies	bake sjokoladekjeks
From English into Norwegian in this particular connection	four weeks long	varer i fire uker
	accept as correct	akseptere som riktig svar
	greater	størst
	unit on geometry	geometriøkt
	"mystery shape"	"den hemmelige figuren"

as Mr. Stone, but rather as Steinar, which could be a Norwegian version of this teacher's first name. If we decide to keep the more formal American setting, most Norwegian teachers would find this different from what they are used to, and they might therefore not experience this as a familiar setting. When making changes from the American names like Mr. Stone and Ms. Wilson to more common Norwegian first names like Steinar and Marianne, we are also adding a potential complexity to the item in that it becomes more difficult to distinguish between the teacher and the pupils in the problem context, since both are referred to by their first names. We therefore had to change some of the items and sometimes include some extra information in order to clearly distinguish between the pupils and the teacher. When making these changes in names, we were conscious about preserving the teacher's gender. When changing into names that are more common in Norway, we tried to find names that were somewhat similar (for example Mr Stone was translated into Steinar because *Stein* means stone in Norwegian) but this was not always done.

Other changes were made in relation to what can be referred to as non-mathematical language. This includes reference to words or contexts that are typical to the American context, but not so familiar in the Norwegian context. In the U.S., for instance, children might be involved in fund-raising. In Norway, children might rather be involved in activities where they collect money and give it to *Redd Barna* – the Norwegian equivalent to *Save the children*. We also do not have school candy

sales in Norwegian schools, so when problems referred to this, we had to change it to the grocery shop. This would be the place where Norwegian children buy candy.

Baking cookies represents a context that is familiar in an American setting, whereas this was not viewed as a familiar activity in an Irish context. In their article, Delaney and colleagues (2008) changed this activity into one of baking scones. Neither of these are familiar activities in a Norwegian setting, so we had to change it into something different. We chose to use the activity of baking chocolate cookies/biscuits, although this is an activity that few Norwegians find familiar. The challenge is to find a good alternative for the translation, and at the same time avoid changing the problem in a way that influences the mathematical challenges that are involved.

The second category of changes relates to the cultural context of the school or the educational system in general. Language used in schools and structural features of the wider educational system are included here. Examples are given in table 2. Since we know the Norwegian school system very well, it was easy to figure out which changes that were necessary. Changes made to the language used in schools are unlikely to compromise the items' ability to measure MKT since these changes do not affect the mathematical substance of the items. But the changes in this category are important to make the item familiar to Norwegian teachers.

The third category relates to the mathematical substance of the items. Examples are given in table 3. We changed units of measurement. In some cases, these translations were straightforward. For example 12 inches might become 12 cm. In these cases the adaptations are similar to context changes such as changes of names and activities (table 1). But not all changes of measurement units were straightforward. In Norway

Table 2. *Examples of school contextual changes made to items*

Type of change	Example from original U. S. form	Example from our translation
School language	practice state mathematics exam	nasjonal prøve
	students' paper	elevarbeidene
	class discussion	fellesdiskusjon
	asks students	ber elever
	write problems	lage oppgaver
Structure of education system	professional development program	etterutdanningskurs for lærere

sweets are not sold in pound, but sometimes in grams. A one-pound bag is for example translated into a 450 gram bag, which is a common size for a bag of chocolate in Norway. This type of change could be more problematic because it risks making the mathematics more difficult for the Norwegian teachers. In this particular item, however, the weight of the bag wasn't used in any calculations.

The category *School mathematical language* refers to changes related to the mathematical language used in schools. The mathematical language used in Norwegian schools of course differs from the language used in schools in the U.S. In most cases precise translations of the terms were possible. But, in Norwegian schools the mathematical language is often translated into a more everyday language. For example, hexagon does have the Norwegian translation *heksagon*, *polygon* could be written the same way in Norwegian as in English and congruent might be translated into the Norwegian word *kongruent*. Our impression is that these more precise mathematical terms are rarely used in Norwegian schools. These terms (more examples are given in table 3) were translated the following way: hexagon – *sekskant*, polygon – *mangekant* and congruent – *helt lik* (English: exactly the same). This could be problematic, because these changes risk making the items easier for the Norwegian teachers.

Representations of mathematical ideas vary from the U.S. to Norway. In Norway for example, comma is used instead of the decimal point, so 1.1 is translated into 1,1. x as multiplication sign is also translated into the sign that is most common to use in Norway: · .

Table 3. *Examples of mathematical changes made to items*

Type of change	Example from original U. S. form	Example from our translation
Units	12 inches	12 cm
	one-pound (bag)	450 grams (pose)
School Mathematical Language	decompose	dele opp
	divide numerators and denominators	deler teller med teller og nevner med nevner
	crossing out	satte strek over
	polygon	mangekant
	congruent	helt like
Representation	1.1	1,1
	.676	0,676
	x (multiplication sign)	· (multiplication sign)

In their presentation of O*ther changes*, Delaney and colleagues (2008) mention changes related to alterations of visual appearance, alternative solutions that were deleted, etc. An example closely connected to representation is presented above. Changing the decimal point to a decimal comma forced us to use space instead of comma to distinguish the decimal numbers.

The use of multiple-choice format for the items is also an aspect worth discussing. This is important because multiple-choice formats has not been widely used in the Norwegian school context and might be unfamiliar to the Norwegian teachers. We have seen indications that this might be changing in Norway, and this appears to be related to the increased use of digital tools in particular. We will discuss this further after the pilot study where one aspect is having teachers comment on the format in a focus group interview. We are prepared to change the format if necessary.

As described above, we have added a category about political correctness. In our process of double translation we had a lot of discussions about the necessity of using a language that was politically correct. In recent curriculum documents and other documents from the Norwegian department, some words have been used in order to describe certain phenomena, and other words have been avoided. One example is the use of the word class, as described above. In order to be politically correct, we have chosen to rewrite the sentences that originally referred to class and use pupils (elever) instead. We could have used group instead, but that could lead to confusion in some instances, because the same word is also used when we refer to group work. For example: "Ms. Wilson's class" is translated into *"Annes elever"* (English: Anne's pupils).

Another example we have chosen to put in this category is use of the verb to learn. In a Norwegian context, we normally refer to learning as the outcome rather than the process. As a result, we find it inappropriate to say that "Mr. Alder's students are learning about ...", since we cannot know if they have actually learned it. In items that refer to the learning process, like in the example just mentioned, we therefore decided to rewrite it somewhat. A Norwegian translation would be: *"Elevene til Anders arbeider med* ..." (in English: "Andy's students are working with ...").

Table 4. *Examples of changes made to items due to political correctness or not*

Type of change	Example from original U. S. form	Example from our translation
Politically correct or not	class	elever
	are learning about	arbeider med

Conclusions

In this article, we have pointed at some issues which indicate that the process of translating and adapting the measures is a highly complex and difficult one. It is not simply a matter of making a literal translation, and several important aspects can be lost in translation. Even changes that appear to be trivial have the potential of making the items more complicated, easier to misunderstand, etc.

We have recently carried out a pre-pilot with five teachers. These teachers have answered the questions in the survey, and participated in a focus group interview directly afterwards. The results from this pre-pilot will be analysed and reported in a forthcoming article, and the aim was mainly to have a quality check of our translation and adaption before the actual pilot study. Although it is somewhat beyond the scope of this article, it might be interesting to point at two preliminary findings from these interviews:

1. The items are perceived as very difficult, even by the experienced teachers.

2. The multiple-choice format is unfamiliar.

If these preliminary findings appear to be relevant for the teachers in the pilot study as well, they might indicate that we are faced with some severe difficulties when we try to use the American MKT measures with Norwegian teachers.

In our pilot study, we are going to evaluate the success of our translations and adaptations by using interviews in addition to analysing the results from the survey. We plan on asking a selection of the teachers whether the items appear authentic to them or not, and whether the mathematical content of the items is of a kind that Norwegian teachers encounter in their regular teaching practice. Data from these interviews might help explaining errors or responses that we would not expect based on previous analysis of the results from the American teachers. These data might also help formulating alternative questions or response options in a future adaption of the items, if necessary. In the Irish study, an analysis of the interview data was fruitful in order to identify which items may cause difficulty for the teachers, and whether the situations and characters described appeared authentic to teachers.

If, when we have carried out and analysed the results from the pilot study, we find that the two issues that have been pointed out above are actual problems, then we might be faced with a situation where we have to acknowledge that using the MKT measures with Norwegian teachers is a dead end. Although this risk of failure is actual and present, we

believe that it is important to try. By going into such a study with a critical view, we might learn something important about the constraints and possibilities that are entangled in the process of translating, adapting and using measures and assessments across language and cultural barriers. The potential rewards from such an endeavour appear to outweigh the risks that are involved, and we find it important for us as researchers to shed light on these issues in order to prevent uncritical use (and abuse) of such measures.

Although our discussion is related to the translation and adaptation of the MKT measures in particular, the main issues we point at should be of significance for researchers involved in translation and adaptation of other kinds of measures and assessments as well. As long as we do not know precisely what has happened in the translation process, and the effects that this has on the results, we should be very cautious when interpreting results from comparative studies like PISA, TIMSS and TEDS-M[2]. We also recommend that researchers to a larger extent discuss issues related to translations in their publications.

References

Adams, R. (2005). *PISA 2003 technical report*. Paris: Organization for Economic Co-operation and Development. Retrieved May 20, 2008 from http://www.oecd.org/dataoecd/49/60/35188570.pdf

Ball, D.L. & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In B. Davis & E. Simmt (Eds.), *Proceedings of the 2002 Annual meeting of the Canadian Mathematics Education Study Group* (pp. 3–14). Edmonton, AB: CMESG/GCEDM.

Ball, D. L., Hill, H. C. & Bass, H. (2005). Knowing mathematics for teaching. Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator*, 29 (3), 14–17, 20–22, 43–46.

Ball, D.L., Thames, M. H. & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59 (5), 389–407.

Begle, E. (1968). Curriculum research in mathematics. In H. Klausmeier & G. O'Hern (Eds.), *Research and development toward the improvement of education* (pp. 44–48). Madisom, WI: Dembar Educational Research Services.

Cooney, T. J. (1999). Conceptualizing teachers' ways of knowing. *Educational Studies in Mathematics*, 38(1–3), 163–187.

Davis, B. & Simmt, E. (2006). Mathematics-for-teaching: an ongoing investigation of the mathematics that teachers (need to) know. *Educational Studies in Mathematics*, 61 (3), 293–319.

Delaney, S. (2008). *Adapting and using U.S. measures to study Irish teachers' mathematical knowledge for teaching.* Unpublished PhD-Thesis, University of Michigan.

Delaney, S., Ball, D., Hill, H., Schilling, S. & Zopf, D. (2008). "Mathematical knowledge for teaching": adapting U.S. measures for use in Ireland. *Journal of Mathematics Teacher Education*, 11 (3), 171–197.

Eisenberg, T. (1977). Begle revisited: Teacher knowledge and student achievement in algebra. *Journal for Research in Mathematics Education*, 8 (3), 216–222.

Falch, T. & Naper, L. R. (2008). *Lærerkompetanse og elevresultater i ungdomsskolen* [SØF-rapport nr. 01/08] . Trondheim: Senter for økonomisk forskning AS. Retrieved December 19, 2008 from http://www.udir.no/upload/Forskning/larerkompetanse_og_elevresultater_i_ungdomsskolen.pdf.

Faulkner, L. R., Benbow, C. P., Ball, D. L., Boykin, A. W., Clements, D. H., et al. (2008). *The final report of the National Mathematics Advisory Panel: U.S. Department of Education*. Retrieved March 26, 2008 from http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf

Geisinger, K. F. (1994). Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6, 304–304.

Grønmo, L. S., Bergem, O. K., Kjærnsli, M., Lie, S. & Turmo, A. (2004). *Hva i all verden har skjedd i realfagene: norske elevers prestasjoner i matematikk og naturfag i TIMSS 2003.* Department of Teacher Education and School Development, University of Oslo.

Haaland, I. & Reikerås, E. (2005). Matematikkfaget ved lærerutdanningen i Stavanger. In M. Lea (Ed.), *Vekst og utvikling. Lærarutdanninga i Stavanger 50 år* (pp. 55–65). University of Stavanger.

Hill, H., Ball, D. L. & Schilling, S. (2008). Unpacking "pedagogical content knowledge": conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39 (4), 372–400.

Hill, H. C., Rowan, B. & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on atudent achievement. *American Educational Research Journal*, 42 (2), 371–406.

Hill, H.C., Sleep, L., Lewis, J.M. & Ball, D.L. (2007). Assessing teachers' mathematical knowledge: What knowledge matters and what evidence counts? In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 111–155). Charlotte, NC: Information Age Publishing.

Johnson, T. P. (1998). Approaches to equivalence in cross-cultural and cross-national survey research. In J.A. Harkness (Ed.), *Cross cultural survey equivalence* (pp. 1–40). Mannheim, Germany: Zentrum für Umfragen, Methoden und Analysen (ZUMA).

Kotsopoulos, D. & Lavigne, S. (2008). Examining "mathematics for teaching" through an analysis of teachers' perceptions of student "learning paths". *International Electronic Journal of Mathematics Education*, 3 (1), 1–23.

KD (2008a). *Kvalitet i skolen. Stortingsmelding nr 31 (2007–2008)*. Oslo: Ministry of Education and Research.

KD (2008b). *Kompetanse for kvalitet. Strategi for videreutdanning av lærere. Midlertidig versjon*. Oslo: Ministry of Education and Research. Retrieved December 19, 2008 from http://www.regjeringen.no/upload/KD/Vedlegg/Grunnskole/Kompetanse%20for%20kvalitet.pdf

Ma, L. (1999). *Knowing and teaching elementary mathematics: teachers' understanding of fundamental mathematics in China and the United States*. Mahwah: Lawrence Erlbaum Associates.

Mullis, I.V.S, Martin, M.O., Gonzales, E.J. & Chrostowski, S.J. (2004). *TIMSS 2003 international mathematics report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.

OECD (2004). *Learning for tomorrow's world: first results from PISA 2003*. Paris: Organisation for Economic Co-operation and Development. Retrieved May 3, 2007 from http://www.oecd.org/dataoecd/1/60/34002216.pdf

Peña, E. D. (2007). Lost in translation: methodological considerations in cross-cultural research. *Child Development*, 78 (4), 1255–1264.

Schoenfeld, A.H. (2007). The complexities of assessing teacher knowledge. *Measurement: Interdisciplinary Research & Perspective*, 5 (2), 198–204.

Shulman, L. S. (1986). Those who understand: knowledge growth in teaching. *Educational Researcher*, 15 (2), 4–14.

Silverman, J. & Thompson, P. (2008). Toward a framework for the development of mathematical knowledge for teaching. *Journal of Mathematics Teacher Education*, 11 (6), 499–511.

Singh, J. (1995). Measurement issues in cross-national research. *Journal of international business studies*, 26 (3), 597–619.

Stiegler, J. W. & Hiebert, J. (1999). *The teaching gap. Best ideas from the world's teachers for improwing education in the classroom*. New York: The Free Press.

Thompson, P. W., Carlson, M. & Silverman, J. (2007). The design of tasks in support of teachers' development of coherent mathematical meanings. *Journal of Mathematics Teacher Education*, 10 (4–6), 415–432.

UFD (2002). *Kvalitetsreformen. Om ny lærerutdanning. Mangfoldig – krevende – relevant. Stortingsmelding nr 16 (2001–2002)*. Oslo: Utdannings- og forskningsdepartementet.

UFD (2005). *Kunnskapsløftet*. Oslo: Utdannings- og forskningsdepartementet.

Notes

1 The items in the actual measures have not been released, and we have therefore used one of the released items as an example here. The Released items can be accessed from http://sitemaker.umich.edu/lmt/files/LMT_Sample_items.pdf

2 TEDS-M includes some MKT items, and the results from this study will be interesting to learn more about. Unfortunately, no results from the Norwegian part of TEDS-M have been published at the time of writing this article.

Reidar Mosvold

Reidar Mosvold is Associate Professor in mathematics education at the University of Stavanger, Norway. His main interest is related to teachers' beliefs and knowledge of mathematics, and their influence on practice.

reidar.mosvold@uis.no

Janne Fauskanger

Janne Fauskanger is Assistant Professor in mathematics education at the University of Stavanger, Norway. Her main interest is related to preschool teachers' mathematical knowledge for teaching and their practice.

janne.fauskanger@uis.no

Arne Jakobsen

Arne Jakobsen is Associate Professor in mathematics at the University of Stavanger, Norway. His interests are mathematics, mathematical knowledge for teaching, and quantitative studies in mathematics education.

arne.jakobsen@uis.no

Kjersti Melhus

Kjersti Melhus is Assistant Professor in mathematics education at the University of Stavanger, Norway. She is especially interested in how to teach mathematics for understanding.

kjersti.melhus@uis.no

Sammendrag

I forbindelse med prosjektet: Learning Mathematics for Teaching (LMT) ble det utviklet måleinstrumenter for å analysere læreres matematiske undervisningskunnskap (MKT). Denne artikkelen presenterer noen av utfordringene som var involvert i et forsøk på å oversette og tilpasse disse målingene for bruk blant norske lærere. Instrumentet ble opprinnelig laget kun med tanke på å bli brukt i en amerikansk kontekst, og en rekke forskjeller mellom de to landene er med på å gjøre dette vanskelig. Vår studie bygger på en tilsvarende irsk studie. I vår studie støtte vi på flere problemstillinger som var tilsvarende de som ble funnet i Irland, men der var også flere nye utfordringer som oppsto når vi forsøkte å oversette og tilpasse måleinstrumentet for bruk i Norge.