ARE WE PLAYING CHINESE WHISPERS?  
ISSUES IN QUESTIONNAIRE DEVELOPMENT

Guri A. Nortvedt\(^1\) and Lovisa Sumpter\(^2\)

\(^1\)University of Oslo, Norway and \(^2\)Stockholm University, Sweden

In this paper, using gender as the topic in focus, we argue that transferring items and surveys from one cultural context to another might be highly problematic. For instance, in the Nordic context, gender issues are addressed in teacher education that reflects how equity is viewed on a societal and political level. Consequently, research on teacher students’ gendered beliefs should acknowledge and take into consideration their knowledge of gender equity during item development. Rather than being translated and adapted, items should be re-constructed and embedded in the context in which they are to be used in order to achieve a valid, reliable instrument. In addition, how gender equity is expressed develops over time, which differs in different cultural contexts. Consequently, time is also a factor to consider.

INTRODUCTION

Both Norway and Sweden have interesting gender patterns regarding mathematics education. In both countries, girls either outperform or perform equally to boys in mathematics, this including upper secondary school. However, in higher education, more men than women opt for mathematically intense programmes, which is a sign of more traditional gender roles (OECD, 2015). Mathematics teacher education, as part of the educational system, also has strong gender patterns in recruitment, with an overrepresentation of female students in kindergarten and primary teacher education programmes (UKÄ, 2016). The result is a segregated workforce (Johnson & Muse, 2016; UKÄ, 2016) and this in two countries with strong traditions regarding equity and equality in both society and education (Imsen, Blossing, & Moos, 2016). Previous research suggests that teachers are one of the main contributors to students’ views on mathematics and other STEM subjects and to their active choice to pursue these subjects in further education (Regan & DeWitt, 2015). Hence, the gender imbalance in the teaching profession might contribute to the gender imbalance in higher education in Norway and Sweden.

Hence, the gender imbalance in the teaching profession might contribute to the gender imbalance in higher education in Norway and Sweden. Researchers have identified a need to understand gender perspectives in mathematics education from a Nordic perspective (Imsen et al., 2016). Therefore, exploring the views of mathematics teacher students’ about gender and mathematics could be key to understanding some of the driving factors of the observed gender patterns. Still, when attempting to reproduce a study by Leder and Forgasz (2010) that has been tested in other countries and on teacher students (Gómez-Chacón, Leder, & Forgasz, 2014), we failed.
After the instrument was adapted and carefully translated into Norwegian and Swedish, the questionnaire was piloted in two Norwegian universities. During piloting, it became apparent that Norwegian student teachers were reacting negatively to some survey items. Although, judging by their response patterns, they found the topic of the survey relevant to some extent, a typical written comment was ‘We are living in a modern world, why do you ask these questions in this way?’ In the words of Andrews and Diego-Mantecón (2015), we realised that we were playing a variation of the game ‘Chinese whispers’, despite adhering to existing good practices for questionnaire translation and adaptation (e.g. Osborn, 2004; Schraw & Olafson, 2015). Beliefs differ considerably from one cultural context to another, making comparative research challenging (Andrews & Diego-Mantecón, 2015; Tuohilampi et al., 2015). However, to better understand how educational beliefs differ culturally valid and reliable research instruments must be developed that legitimise both the act of comparison and the categories used (Clarke, 2013).

The purpose of this research report is to discuss criteria and best practices in developing culturally valid questionnaires that enable investigation of student teachers’ gendered beliefs about mathematics and mathematics education in more than one educational context.

CULTURE, GENDER AND THE NORDIC MODEL

Previous research has identified several issues regarding cross-cultural studies in mathematics education (Andrews & Diego-Mantecón, 2015; Clarke, 2013). A major issue is how to define culture. One solution, although somewhat narrow, is to view cultures as national contexts (e.g. Andrews and Diego-Mantecón, 2015). A more general approach can be found in Clarke (2013), who discusses the cultural validity of cross-cultural comparisons in mathematics education. Although not distinguishing the various levels of how a culture can be operationalised, his definition enables a view of culture as functioning as a lens that is not restrained to national borders. Here, we follow Clarke’s (2013) definition. Furthermore, we argue that within a national context, several particular cultural, that is, sub-cultural, practices and identities exist that must be taken into account when conducting cross-cultural research.

The Nordic education model can serve as an example of a specific cross-cultural context, spanning several countries, including Norway and Sweden. It has been identified with attributes shared by the various countries regarding equal access to education, including no segregation by abilities, social class or gender (Imsen et al., 2016). Here, we focus on the latter aspect, in which gender refers to ‘feminine and masculine, characteristic and culture dependent traits attributed by society to men and women’ (Wedege, 2007, p. 252). A main trait of the Nordic education model is the over-arching perspective regarding gender equity and equality: both Norway and Sweden have implemented national laws and policies securing equal rights in terms of education (Hedlin, 2013; Imsen et al., 2016). This Nordic perspective has
traditionally been reflected in the national curricula: for instance, in Sweden it was clearly stated in 1962 that education in all subjects should be given to every child as a legal right, based on the ideas of equity and equality (Imsen et al., 2016). While equal rights are explicit in policy documents, students and teachers at various levels express gendered views about mathematics and mathematics education (OECD, 2013; Sumpter, 2012; Sumpter, 2016). At the same time, many teachers (in Sweden) believe that gender issues are not relevant to their teaching, indicating that the ‘gender problem’ is solved (Gannerud, 2009). This illustrates a tension among various driving factors: on one hand the observed structural gender and symbolism at the individual and group levels, and on the other hand the constituted societal and political views.

WHERE WE WENT WRONG – THE PILOT STUDY

This study started out with the intention to replicate the study by Leder and Forgasz (2010) utilizing a tool previously applied in other studies involving various kinds of samples and cultures (e.g. Forgasz et al., 2014; Gómez-Chacón et al., 2014). As a first step, the survey was translated from English into Norwegian and Swedish individually by the two authors of this paper. To validate the translation, both translations were compared and back-translated, following the procedures presented in Andrews and Diego-Mantecón (2015). In addition, between-language comparisons were conducted. Prior to translation, some adaptions were made. For instance, as the target population was prospective teachers, background questions aimed at identifying teacher-education programme and year of study were added. We also added a question about whom the respondents believed to be best suited to teach mathematics, a female or a male teacher. The question, ‘Should students study mathematics when it is no longer compulsory’ was removed, as this is not relevant in a Norwegian and Swedish contexts, in which mathematics is compulsory and all students must take mathematics courses in upper secondary school. In addition, some items were carefully adapted to better conform to Norwegian and Swedish ways of expression. For example, the question ‘Who are more suited to being scientists, boys or girls?’ was changed to ‘Who are more suited to work in professions in which you apply mathematics, e.g. engineer and chemist?’.

Next, students in two Norwegian universities were invited to participate in the pilot study, in accordance with good practice for questionnaire development (Andrews & Diego-Mantecón, 2015; Schraw & Olafson, 2015). The responses indicated that some questionnaire items were culturally skewed, including the item ‘Who is better at mathematics, boys or girls?’ The majority of students (56%) stated that boys and girls are equally good, and 28% stated that boys are better. However, comments indicated that some students experienced a dilemma when replying to this item. One male student commented:
When we speak about ‘good at mathematics’ I am thinking that you think of this related to school outcomes (not the ability to assimilate knowledge and problem solving skills). And then, I believe that previously boys have been more interested in mathematics and science programs in higher education, but that now more and more girls study mathematics. And I believe that if someone works with something (for instance mathematics) they can be really good at it.

Based on this and other comments, including expressions that the survey questions were inappropriate because we are ‘living in a modern world’, we suspected that the students in the pilot study were experiencing the tensions between a local culture and the constituted national context of gender. Simply put, we were asking the wrong questions if we wanted to study prospective teachers’ views about gender and mathematics education. Furthermore, the survey did not allow the participants to demonstrate knowledge of how gender can be viewed in various cultural contexts, both within and between different groups. Hence, to continue the work of adapting the survey, we needed to address various cross-cultural research criteria.

**WHAT IS MISSING? CULTURALLY AWARE CRITERIA**

The steps for adapting questionnaires presented by Andrews and Diego-Mantecón (2015) involve the following four principles taken from Osborn (2004): conceptual equivalence, measurement equivalence, linguistic equivalence and sampling. Measurement equivalence and sampling apply only to the measurement phase and therefore are not instrumental to addressing the cultural aspects of questionnaire development. Consequently, these principles will not be discussed further. Linguistic equivalence can be achieved through good translation practices (Andrews & Diego-Mantecón, 2015), hence leaving us with the challenge of constructing conceptual equivalence. The goal of conceptual equivalence is to ‘provide conceptual definitions that have equivalent, though not necessarily identical meaning in various cultures’ (Osborn, 2004, p. 269). However, Osborn (2004) does not discuss what it means to be equivalent while not identical, or how researchers can achieve conceptual equivalence other than by using an inside/outside perspective as part of the development. A potential inference is that what is missing is a more nuanced understanding of what cultural awareness could constitute. Therefore, we turned to Clarke (2013), who uses seven dilemmas when discussing how international comparative research using questionnaires might be undertaken: (1) cultural-specificity of cross-cultural codes; (2) inclusive vs distinctive; (3) evaluative criteria; (4) form vs function; (5) linguistic preclusion; (6) omission; and, (7) disconnection. These dilemmas might be used to develop the needed culturally aware criteria. In this discussion, we will focus on six of these dilemmas. Form vs function mainly concerns classroom or teaching activities that could be interpreted as having the same form but different functions. As our study does not focus on activities, this dilemma is less relevant than the others are.

When creating questionnaires, something inevitably will be omitted. This is the sixth of Clarke’s (2013) dilemmas: when research is unable to capture everything that is or
needs to be observed. When we asked prospective teachers about their gendered views, their written comments suggested that we had indeed omitted changes in the construct that have taken place in society over time. Item 6 from the original study illustrates this dilemma. The item addressed the question of whether girls or boys are better at mathematics. To elicit more information about the students’ beliefs, this item was accompanied by a follow-up question: ‘Has this changed over time?’ Previous research has concluded that gender is indeed a construct that has been changed over time (Hedlin, 2013; Imsen et al., 2016). Student responses to this and other follow-up question indicated that some students acknowledged that changes have taken place and some found the very question(s) outdated. From the students’ comments, we concluded that not only did we need to formulate items differently, but also that follow-up questions needed to be added to several more questions in future versions of the questionnaire. This, for instance, could enable us to distinguish between groups who believe that gender differences have never existed and those who think that there has been development (c.f. Sumpter, 2016). Besides allowing us to distinguish between these subgroups, we might also be able to generate nuances in views with time as a factor.

Clarke’s (2013) fifth dilemma, linguistic preclusion, is addressed somewhat by Osborn’s (2004) linguistic equivalence and also is closely linked to Clarke’s first two dilemmas, cultural-specificity of cross-cultural codes and inclusive vs distinctive. In Norwegian, ‘kjønn’ is used to illustrate gender, while in Sweden, separate concepts exist to distinguish gender and sex (genus and kön). In the questionnaire, all three terms are omitted, and ‘boys’ and ‘girls’ and ‘male’ and ‘female’ are used consistently to address Clarke’s evaluative criteria and the linguistic preclusion criteria. In Norwegian and Swedish, the chosen terms, ‘boys’ and ‘girls’, can hold all necessary meanings/interpretations. As stated above, this is related to the first of Clarke’s (2013) dilemmas, cultural-specificity of cross-cultural codes. When applying concepts across cultures, awareness of potentially different connotations is crucial. Our experiences with the first version of Item 6, asking whether girls or boys are better at mathematics, illustrates this: the question was relevant in previous cross-cultural studies (Forgasz, Leder, & Tan, 2014; Gómez-Chacón, Leder, & Forgasz, 2014), and possibly to some of the participants in our pilot, judging by the response patterns observed. Still, to others, this question represented an ambiguity that positioned them between the culture of the Nordic model and the observed gender patterns regarding school outcomes and recruitment into higher education, as theorised in the Introduction. This ambiguity applies to the second dilemma, inclusiveness vs distinctiveness, as well. Sacrificing distinctive (cultural) details results in questionnaire items that respondents regard as too open and/or too general. Hence, we suggest that providing a contextual frame could resolve this issue for our survey, since this frame could address the tension between the Nordic model and observed practices. Therefore, our solution is that instead of a series of independent, multiple-choice questions as used in Leder and Forgasz (2010), a series of items
linked to the same contextual frame could allow respondents a culturally acceptable interpretation span (see Figure 1).

![Figure 1. Item 1 of the revised questionnaire.](image)

In addition, different cultures value different concepts differently. This most likely also applies to various groups within each culture and, hence, in the present study, it could be addressed by reformulating the original item into the form in Figure 1. This reformulation enables us to address various interpretations of the topic addressed and at the same time allow students participating in the study to show by their responses their interpretations of the views of the groups listed. In addition, students could be given a follow-up question that might ask them to reflect on changes over time for one or more of the listed groups.

Regarding the seventh and last dilemma, which is misrepresentation through disconnection, any survey targeting individuals’ beliefs or views will be disconnected from the context being asked about (c.f. Tuohilampi et al., 2015). Although not a perfect remedy, providing a short introduction that aims to provide a cultural context might address this dilemma to some extent.

**CONCLUDING REMARKS, INCLUDING IMPLICATIONS**

Our failure to reproduce research undertaken in other cultures brought the issue of cultural validity to the surface. Through our discussions, we have sought to elaborate on this issue from various perspectives. Many of the established good practices cannot alone provide the awareness necessary when conducting cross-cultural research, at least when we speak about phenomena that are socially constructed and culturally/context conditional, as is gender. To avoid accidentally playing ‘Chinese
whispers’, besides focusing on content and construct validity, we also need to consider a cultural validity in which culture encompasses more than just national borders (c.f. Andrews & Diego-Mantecón, 2015). We argue that within a national culture, several ‘cultures’ exist, and it is necessary to shape questionnaires in such a way that participants can interpret the construct in similar ways, as much as that is possible. Our solution is (1) to provide framing that explains the cultural setting in which the questions are posed to enable interaction with the items and (2) to offer multi-dimensional choices that allow flexibility in the responses. Another possible implication is that our solution for constructing questionnaires in cross-cultural research is applicable not only to beliefs about gender but also to other beliefs or other affect research focusing on individuals’ conceptions. We make this inference in reference to the conclusions of Tuohilampi et al. (2015), who argue that you must understand the construct in focus from the specific contexts and ‘that theoretical affective models should be considered with respect to a given culture’ (Tuohilampi et al., 2015, p. 1644). Based on our failed attempt, we could not agree less.

References


