THE LANGUAGE BARRIER: USING CONCEPT INVENTORIES TO TEACH SCIENCE TO FIRST-YEAR UNIVERSITY STUDENTS

Jurgen Schulte[a], Neela Griffiths[b]

Presenting Author: Jurgen Schulte (jurgen.schulte@uts.edu.au), Neela Griffiths (Neela.Griffiths@uts.edu.au)
[a]School of Physics and Advanced Materials, University of Technology Sydney, Ultimo NSW 2007, Australia
[b]Academic Language and Learning Group, Institute of Interactive Media and Learning, University of Technology Sydney, Ultimo NSW 2007, Australia

KEYWORDS: transition pedagogy, discipline-specific language in teaching, language in tests

ABSTRACT
The effective use of language in science is a perennial topic in science education. In studying science students are confronted not only with the scientific concepts of the disciplinary area but also with its discipline specific language. In English, this challenge is compounded by the overlap in meaning between discipline specific terminology and everyday language. Science concept testing tools, which have been developed to deal with these issues, are widely used for testing understanding of science concepts. However, Science educators are facing new challenges with the increasing diversity of the student cohort accessing tertiary education. These students' educational, cultural and linguistic backgrounds combined with their unfamiliarity with the academic discourse of their discipline may be an obstacle to their smooth transition. In the past, the specially written Concept Inventories supported lecturers and students in this transition but these inventories were not designed for the student cohorts of the twenty-first century. This paper reports on our study of one of the more proven concept testing tools 'The Force Concept Inventory' (physics) and its application in large first-year classes with non-traditional students.

INTRODUCTION
The intentional design of learning, teaching and assessment approaches forms an integral part of any successful transition pedagogy (Kift, 2009). Our research focuses on domain specific language barriers that students face in their first year at university and how these may affect students from low socio-economic (LSES) and non-English speaking backgrounds* (NESB). Here, we present preliminary results of our study on the language used in assessment and testing of students' comprehension of concepts in physics using a well established physics force concept test. While our study is being conducted within the domain of the teaching and learning of physics, we believe that the observations made in this study may inform learning, teaching and assessment approaches in other disciplines.

Concept inventories are common instruments in science education. They are used to evaluate students' learning of discipline specific concepts in science as well as the effectiveness of teaching instructions. The Force Concept Inventory (FCI) is a multiple-choice test designed to assess student understanding of concepts in Newtonian mechanics (Hestenes & Halloun, 1995). One concern with concept inventory tests in science is the question of whether they are a reliable measure of the understanding of scientific concepts and scientific literacy or whether they are more likely to reflect a student's language proficiency. This is a particular concern in physics where there is a considerable overlap of common language words (for example, heat, force, work), which in a physics context often take on quite different meanings and even change their ontological position within the language without any semiotic indication. For a non-English speaking background student (NESB) such words have to be compartmentalised in the context of the related physics concept. Some of these concepts are expressed metaphorically as they are much more difficult to conceptualise (for example, heat, field, wave function).

In the early 1990s, when performance assessments in science started to increase considerably in the U.S. educational system, it was found that performance outcomes continued to lag behind expectations. This trend was attributed to the rapid increase in English language learners (ELLs, equivalent to the term NESB used in Australia) entering the U.S. educational system. Hence, the question arose whether science tests and evaluations were biased toward linguistic modes of comprehension and alternative strategies of testing ought to be developed. The introduction of so-

called sheltered English science classes was one of many strategies to address language and culture related learning difficulties. (Sheltered English science classes only comprise students of non-English speaking background and the teacher as well as students may use languages other than English in class). With the prospect of ELLs comprising one-fourth of the American student body by 2026 (Garcia, 2002), it was found that performance affecting language and cultural biases in science performance tests as well as in various modes of teaching needs closer attention in order to maintain a high level of science literacy. The problem is a complex one as highlighted by Shaw (1997):

> Thus, there is no simple answer to the question whether performance assessment for ELLs (English language learners) measures English language proficiency or scientific literacy. Even within the same carefully designed assessment, different questions are influenced by different abilities.

Among the more critical types of assessment are exams. Bosher and Bowles (2008) found some evidence that language may be a source of construct-irrelevant variance for students with non-English speaking backgrounds when they take exams especially those based around multiple-choice questions (MCQs), which are frequently used in large first year classes.

Wellington and Osborne (2001) in their seminal book on language and literacy in science education argued that among other things science teachers are language teachers. Although their reference was with respect to science education at primary and secondary school level, the situation is not very different for a lecturer teaching first year physics to non-science majors (in our case engineering students). In an Australian context, first-year physics classes are moderately large, ranging from about 200 to 600 students from a multitude of non-English speaking backgrounds and associated cultural origins. Typically, one would find about 40-60% of students fall into this category, speaking 30 to 40 distinctively different languages at home. This does not leave much room for the lecturer to clarify concept specific words with short phrases in non-English language since there is no clear language majority among students (as is the case of an equivalent American class where some use of Spanish, for example, can support a large non-English speaking cohort).

In a recent study, Moore (2007) identified students' 'culturally tainted' language as an obstacle in science education and learning. Unlike in the studies involving clearly non-English speaking students as mentioned above, Moore compared performance and progress of native English speaking students of African-American ethnicity (low socio-economic background) with their white English speaking peers. This study has relevance in an Australian context since over the past few years, policy decisions have required Australian universities to increase their intake of LSES students considerably. Moore noted that for students of such backgrounds it is difficult to cross language and cultural boundaries to obtain the language and culture of power exercised within academic settings:

> Teaching the nature of science more explicitly and incorporating students' home cultures and languages at the same time has a great potential for opening the gate of language and discourse in the science class room (Moore, 2007).

Brookes (2006) examined the linguistic aspects of concepts in physics more closely, in particular the role of systematic 194ypothesized of language in the process of teaching and learning of concepts. In Brookes' theoretical framework of grammatical ontology, it is shown that a systemic incorrect use of language in physics may be responsible for students' learning difficulties, development of misconceptions as well as the continuing reinforcement of misconceptions. What we mean by systemic incorrect use in this context is that the spoken as well as written language widely used in physics by teachers, lecturers and academics and in common English language physics text books world-wide (secondary school, university) all show similar misuse or incorrect use of grammatical ontology. For example, a sentence with correct functional grammar and ontology is (Brookes, 2006)

> "John [agent] kicked [process] the ball [medium]". Here "John" and "the ball" are grammatical participants, functioning grammatically as objects or matter. "John" and "the ball" are naturally 194ypothesize as matter in some sense. Thus the grammatical ontology and lexical ontology match. In contrast, in "heat [medium] flows [process] from the environment to the gas" the grammatical ontology and lexical ontology do not match. In this sentence a physicist would recognize heat to define a process of movement of energy into the system (lexical ontology). But

grammatically "heat" is functioning as a participant, namely heat is the matter that is flowing. The grammatical function of the term "heat" and the lexical ontology of "heat" contradict each other."

The sentence suggests that the flow of heat is a material process, which in a physics context is fundamentally incorrect and triggers subsequent incorrect logical assumptions. One could say that this phrase is of a more metaphorical nature rather than representing a factually correct sentence. Brookes (2008) refers to such constructs as grammatical metaphors. There ought to be an easier way for students to learn physics than by wading through a cloud of metaphors. Force is another prominent word in physics that is often used in a grammatically metaphoric way, making it rather difficult for students to decipher 'force' metaphors and to acquire a good understanding of the concept of 'force' in physics.

LINGUISTIC MODIFICATION OF THE FORCE CONCEPT INVENTORY (FCI)
The decision to modify the widely used Force Concept Inventory (FCI) by Hestenes, Wells and Schachhamer (2006) was based on the confusion that the complex concept of force causes. Traditionally, as Brookes alludes to, students' difficulties with the concept of 'force' have been attributed generally to student misconceptions, preconceptions or alternative conceptions (Hestenes et al., 2006). However, Brookes' argument is that although students may arrive at university with ineffective epistemological resources these do not account for all the students' difficulties with the concept. He suggests that it is the way that physicists represent the concept through their use of language, which directly influences the students' reasoning rather than the students' belief systems (Hestenes et al., 2006). This could be an additional challenge for NESB students, whose belief systems may be influenced by their first languages in which the concept of 'force' might be interpreted differently. Moreover, the apparent familiarity of the word 'force' due to its everyday meanings may also cause confusion and misunderstanding for students without a physics background.

A further challenge for NESB students is that exams not only test their discipline specific knowledge but also their English language proficiency and cultural awareness. In an effort to rectify this, researchers in the United States have linguistically modified a range of exams to assess whether this process would improve the results of NESB students without affecting the results of native-speaker students (Bosher & Bowles, 2008). Linguistic modification is a process characterised by maintaining the 'key content area vocabulary and terms' and reducing the 'semantic and syntactic complexity of the English' (Bosher & Bowles, 2008). Findings from these studies, although exploratory, suggest that the process has increased NESB students' comprehensibility of the tests without impacting on other students' results.

METHOD
In 2013, the authors, a physics lecturer and an academic language and learning developer (linguist), received a university grant supporting first-year students with transition. One objective of the grant was to design a series of language focussed interventions for a compulsory first-year Physics subject for Engineering students. These students are from diverse educational, cultural and linguistic backgrounds and many are NESB students. Due to the problematic nature of the language of 'force' in physics, the linguistic modification of the FCI was included in our intervention. The first step in the modification was an independent reading of the FCI with the intention of identifying linguistically or culturally problematic questions. At our follow-up meeting, we agreed that more than half the questions in the original version were potentially challenging for our students particularly as the physics lecturer was required to clarify ambiguities in a number of the questions. The next step was to produce a version which would be more accessible to our student body. In order to do this, we identified and prioritised a number of common threads of possible difficulty. These included questions which were: syntactically dense; semantically ambiguous; wordy; had weak clarity of expression; and/or, were culturally bound. Changes would be applied to the question's stem and/or hypothesize as required. A further modification was the decision to replace the verb 'exert' as it is difficult to translate precisely into other languages. As noted above, exams in English for NESB students can seem like English reading tests; however, despite the fact that the average item-response time for NESB students is longer (Bosher & Bowles, 2008) we did not consciously try to shorten the 195ypothesize rather we tackled their lexical complexity.

Some examples of the linguistic modifications made to the question stems are below. The intention was to make the question rubric brief and concise. For example,

Question 5:
Original: 'The accompanying figure shows a . . . '
Modified: 'Figure 1 shows a . . .'

We also replaced some of the culturally bound expressions, for example,

Question 8:
Original: 'The figure depicts a *hockey puck* sliding . . . ' (Italics added)
Modified: 'Figure 1 shows an *object* sliding . . .' (Italics added)

One example of a modification which combined both of the above, is,

Question 14:
Original: 'A bowling ball accidentally falls out of the cargo bay of an airliner as it flies along in a horizontal direction.'
Modified: 'A care package accidentally falls out of an airplane as the airplane flies along in a horizontal direction.'

We also identified and replaced rather out-moded terminology such as:

Question 10:
Original: 'increases for a while and decreases thereafter.'
Modified: 'increases for a while and decreases after that.'

Once the modification process was complete both the original and modified versions of the FCI were uploaded to a response system. One group of students (n= 71 respondees) was given the modified inventory and the other group (n=284 respondees) the original. The two groups of students belonged to the same class and had the same lecturer but attended lectures at different times due to their individual timetable constraints. Each group included a diverse range of language backgrounds. The FCI test was administered after the 'force' concept had been covered in class and tested in an exam (mid-semester).

ANALYSIS
In this section we will make some general comments about the students' answers to both versions of the FCI. For brevity, we have selected two examples. Both received a higher number of correct responses on the modified version. One example has a modified question stem and in the other the verb 'exert' is replaced. We will also comment on a question which reflects the students conceptual confusion with the meaning of force.

Example 1:
The description rubric for Question 20 was rewritten from:

Original:
'The positions of two blocks at successive 0.20-second time intervals are represented by the numbered squares in the figure below. The blocks are moving toward the right.'
Modified:
'Two blocks are moving to the right along a position axis. Their positions are shown at successive 0.20s time intervals from 1 up to 7.'
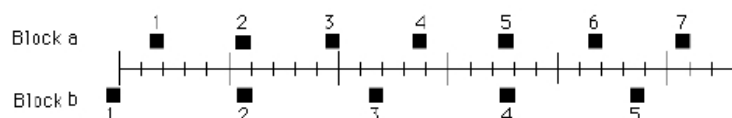
No changes were made to the 196ypothesize. In Question 20, 77% selected the correct answer in the modified version compared to 68% in the original. The first sentence in the original version includes a long, densely packed and confusing noun phrase as well as an unnecessary prepositional phrase. These were removed and replaced with a first sentence, which clearly links the position and the direction of the blocks directly to the diagram. The inclusion of the phrase 'time intervals from 1 up to

7' has the same function. Overall, the descriptive rubric is simpler, more concise and meaningful for the students.

Example 2:
On balance, and this was not unexpected, both groups experienced more difficulty with the questions which had the longer 197ypothesize especially in instances where the 197ypothesize were very similar, were lexically dense and/or used comparative structures. Question 4, a question which we had modified quite considerably, was one of the exceptions to this as the group completing the modified version did better on it despite the length and similarity of the 197ypothesize and the use of comparative structures. It also used the verb 'exert' in each of the original 197ypothesize. The original and modified versions are below.

Question 4:
Original:
A large truck collides head-on with a small compact car. During the collision:
a)   the truck exerts a greater amount of force on the car than the car exerts on the truck.
b)   the car exerts a greater amount of force on the truck than the truck exerts on the car.
c)   neither exerts a force on the other, the car gets smashed simply because it gets in the way of the truck.
d)   the truck exerts a force on the car but the car does not exert a force on the truck.
e)   the truck exerts the same amount of force on the car as the car exerts on the truck.

Modified:
A large truck collides head-on with a small car. During the collision:
a)   the force from the truck onto the car is greater than the force from the car onto the truck.
b)   the force from the car onto the truck is greater than the force from the truck onto the car.
c)   neither applies a force onto the other, the car gets damaged simply because it gets in the way of the truck.
d)   there is force onto the car from the truck but there is no force onto the truck from the car.
e)   there is the same amount of force from the truck onto the car as there is from the car onto the truck.

Although a higher percentage answered the modified version correctly (response e, 76% vs. 71%), it is apparent from this result that, overall, students' understanding of the concept is still unclear.

For the most part, the students did better in both versions on the questions which required less reading time due to shorter question stems, 197ypothesize or which used diagrammatic representations. However, we observed three exceptions to this. Questions 5, 18, and 30 caused some level of confusion in both versions of the FCI. It can be assumed that this is partly due to the 'force' concept itself but it could also be due to the way that the term 'force' is used in these questions. Research has shown that some of the confusion students experience in relation to the force concept could be due to the way that physicists represent the concept through the language which they use (Brookes, 2006). In physics textbooks, force is used as the medium of the sentence or as the agent. Indeed, 'The way in which physicists use "force" grammatically, and the many associated conceptual metaphors, seemed to contradict each other' (Brookes, 2006). An analysis of the ways in which 'force' is used in these questions reflects these contradictions. Clearly, we will need to represent the 'force' concept more consistently in future iterations of the modified version.

The original and modified versions of Question 5 are below (the 197ypothesize in Questions 18 and 30 are similar).

Question 5:
Original:
Consider the following distinct forces:
1. A downward force of gravity.
2. A force exerted by the channel pointing from q to O.
3. A force in the direction of motion.
4. A force pointing from O to q.
Which of the above forces is (are) acting on the ball when it is within the frictionless channel at position "q"?

Modified:

Consider the following distinct forces:

1. A downward force of gravity.
2. A force coming from the channel pointing in the direction from q to O.
3. A force in the direction of motion.
4. A force pointing from O to q.

Which of the above forces is (are) acting on the ball when it is within the frictionless channel at position "q"?

CONCLUSION

Our preliminary findings on the impact of the language in physics tests on students' understanding of concepts and test performance show that test outcomes do not necessarily represent an indication of students' actual learning or understanding of physics concepts. The widespread practice of using key concept words such as 'force' in a language of ontological metaphors in introductory physics textbooks, over 80% in the books by Cutnell and Johnson, and by Halliday, Resnick and Walker for example (Brookes, 2006), may also be an indication of how widespread the use of incorrect ontological grammar is in lectures and assessable tests. Our findings support the idea that the inconsistent and conflictual use of conceptual language in physics may be responsible for students' learning difficulties as 198ypothesized in Brookes' theoretical framework of grammatical ontology. Perhaps it is the metaphorical nature of the language used in the physical sciences that contributes to turning away many students from considering science as a choice. In addition to this, generally speaking, science students have not been prepared for the amount of reading of technical texts which is expected of them at a tertiary level. Thus, it is not surprising that the students' apparent difficulties with reading comprehensibility in the FCI supports research conducted in the U.S.

Grammar ontology originated comprehension difficulties in physics (or generally in the physical sciences) is an area in transition pedagogy that has not been well studied. Grammar ontology originated comprehension difficulties are not only adding to the many difficulties that students encounter during transition to university, they may also have a long lasting impact with respect to student learning of physics concepts and students' performance in standard tests and exams. This becomes a particular issue for NESB students who at this stage are still struggling with 'normal' English grammar and technical vocabulary. An extended study including pre- and post learning evaluations of our non-science majors' first year physics classes using the "language modified" FCI test is underway and will shed more light into the impact of this language barrier. Further results including the complete, modified FCI test will be published elsewhere.

ACKNOWLEDGEMENTS

REFERENCES

Bosher, S. & Bowles M. (2008). The effects of linguistic modification on ESL students' comprehension of nursing course test items. *Nursing Education Research, 29*(3), 165-171.

Brookes, D. T. (2006). *The Role of Language in Learning Physics.* Rutgers Thesis. The State University of New Jersey, USA.

Garcia, E. (2002). *Student cultural diversity: Understanding and meeting the challenge (*3rd.ed.). New York, NY: Houghton Mifflin Company.

Hestenes, D. & Halloun, I. (1995). Interpreting the Force Concept Inventory. *The Physics Teacher, 33*, 502-506.

Hestenes, D., Wells, M., & Swachhamer, G. (1992). Force Concept Inventory. *The Physics Teacher, 30*, 141-158.

Kift, S. (2009). *Articulating a transition pedagogy to scaffold and to enhance the first year student learning experience in Australian higher education.* Final Report for ALTC Senior Fellowship Program 2009. Retrieved August 5, 2013, from http://www.olt.gov.au/resource-first-year-learning-experience-kift-2009.

Moore, F. M. (2007). Language in science education as a gatekeeper to learning, teaching and professional development, *Journal of Science Teacher Education*, *18*, 319-343.

Shaw, J.M. (1997). Threats of the validity of science performance assessments for English language learners, *Journal of Research in Science Teaching*, *34*(7), 721-743.

Wellington, J. & Osborne, J. (2001). *Language and literacy in science education*. Buckingham, UK: Open University Press.