

**This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.**

Author(s): Arffman, Inga

Title: Unwanted literal translation : An underdiscussed problem in international achievement studies

Year: 2012

Version:

Please cite the original version:

Arffman, I. (2012). Unwanted literal translation : An underdiscussed problem in international achievement studies. Education Research International, 2012, Article 503824. <https://doi.org/10.1155/2012/503824>

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Review Article

Unwanted Literal Translation: An Underdiscussed Problem in International Achievement Studies

Inga Arffman

Finnish Institute for Educational Research, University of Jyväskylä, P.O. Box 35, 40014 Jyväskylä, Finland

Correspondence should be addressed to Inga Arffman, inga.arffman@jyu.fi

Received 29 March 2012; Accepted 18 June 2012

Academic Editor: Gwo-Jen Hwang

Copyright © 2012 Inga Arffman. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In international achievement studies, a common test is typically used which is translated into the languages of the participating countries. For the test to be valid, all the translations and different-language test versions need to be equally difficult to read and answer. An underestimated and underdiscussed threat to this validity is unwanted literal translation. This paper discusses the problem of unwanted literal translation in international achievement studies. It defines what is meant by unwanted literal translation and explains why it is a threat to the validity of international achievement studies and why it is so difficult to avoid. It also discusses problems there have been when translating these tests which may have promoted unwanted literal translation and provides suggestions on how to improve the translation practices so as to ensure that the translations are in as natural and idiomatic language as possible.

1. Introduction

Recent years have witnessed a huge interest in international achievement studies, whose results are increasingly used, for example, in educational decision making. Studies have been conducted, for example, by the International Association for the Evaluation of Educational Achievement (IEA), the Educational Testing Service (ETS), Statistics Canada (STATCAN), the Organization for Economic Co-operation and Development (OECD), the United Nations Educational, Scientific and Cultural Organization (UNESCO), and the Southern Africa Consortium for Monitoring Education Quality (SACMEQ). In all these studies, a common test has been used which has been translated or adapted into the languages of the participating countries. (In this paper, the term “translation” refers to the process of reproducing a text originating in one language and culture for use in another language and culture. The term thus covers all kinds of between-language meaning transfer, from close translation to adaptation, or the making of changes to the target version so as to make it more suitable for the target population (e.g., changes in currency or measurement units). The term “adaptation,” accordingly, refers to a special subtype of translation.)

When translating large-scale international achievement tests, the demands on the translations are extremely high, much more so than in any other type of cross-cultural comparative studies [1, page 49]. This is because, contrary to what is often the case in other cross-cultural studies, in international achievement studies the different-language versions need to be equivalent or comparable, not only in meaning but also in difficulty. Besides, unlike other cross-cultural studies, in international achievement studies the instruments contain not only items but also stimulus texts, which also need to be equivalent in meaning and difficulty. If the materials are not equivalent, valid comparisons between countries are not possible.

From a cognitive point of view, the requirement for equivalence in difficulty means that the mental effort required of testees to respond to the items needs to remain the same across languages. No version must place a heavier cognitive load [2], [3, pages 93-4], or consume more of the limited processing capacity of testees’ working memory compared to the other versions. This, in turn, requires, among other things, that all different-language items and stimulus texts be equally easy to read and understand. If this is not the case, if some items or stimuli are harder to understand

TABLE 1: Examples of unwanted literal translations and idiomatic translations.

Unwanted literal translation	Idiomatic translation
(1) On language and image transmission on the one hand, respectively data transmission on the other hand, different requirements are set regarding the transmission quality to be observed.	The quality requirements for voice and video transmission are different from those that apply to data transmission.
(2) He then his mishap reported to the police, who are the thief intensively searching.	He then reported his misfortune to the police, who are searching diligently for the thief.
(3) Madame Odette, passenger with destination Douala, is demanded on the telephone.	Ms. Odette, passenger for Douala, you are wanted on the phone.

than the others, more working memory is needed to decode and make meaning of them—reading and comprehension being the first phase in the survey response process—and less memory is left for actually responding to the items (i.e., retrieving the relevant information, forming the judgment, and editing the answer, e.g., [4, page 8]). Readers of these versions would then be at a disadvantage, which, in turn, would jeopardize the validity of inferences made on the basis of the test. An underestimated and underdiscussed threat to the validity of international achievement studies is unwanted literal translation or unwanted literal rendering.

2. Purpose and Outline of This Paper

This paper discusses the problem of unwanted literal translation in international achievement studies. The purpose is twofold: to increase awareness of the threat unwanted literal translation poses to the validity of these studies and to discuss problems there have been when translating these tests which may have endangered their idiomaticity and to provide suggestions on how to ensure that the translations are in as idiomatic language as possible.

The paper first defines unwanted literal translation and explains why it is a threat to the validity of international achievement studies and why it is so difficult to avoid. Then, after briefly describing the translation procedures and practices in international achievement studies, it discusses factors that have an impact on how literally texts are translated and how these factors can be improved so that the ensuing translations would be as idiomatic as possible. The paper finishes with a summary of the lessons learned and suggestions for future studies.

Theoretically, the discussion is grounded in (cognitive) Translation Studies, whose principles should guide all translation work but often seem to have been forgotten in test translation [5, page 118]. When discussing translation procedures and practices in international achievement studies, the focus will be on studies conducted by the OECD and IEA, because for other studies very little, if any, of such data is available. However, since the cognitive processes and principles of translation are the same in all translation contexts and since naturalness and idiomaticity are goals in most translation contexts, the discussion is believed to be helpful also when translating other cross-national achievement tests and even other types of cross-national tests. By the same token, even though the paper

focuses on unwanted literal translation—because much less attention has been paid to it than to other types of translation problems—the suggestions for improvement will be helpful also when solving these other problems.

3. Definition of Unwanted Literal Translation

Unwanted literal translation, or interference, is an extremely common problem in translation. As used in this paper, it refers to translations that are rendered word for word and strive to stay formally (e.g., lexically and syntactically) as close to the source text as possible (cf. [6, page 159] [7, page 208]), with the result that the target text becomes odd, unnatural, and cumbersome. (Not all literal translations may be described as “unwanted”. e.g., often, especially when translating between closely related languages, literal renderings are completely idiomatic and natural. Also, literal translations are frequently used intentionally, in linguistic contexts, to clarify the syntax and structure of a foreign and often exotic language.) Unwanted literal translation can also be seen as an opposite to idiomatic translation, which, for its part, refers to translations that attempt to read like normal and authentic target language texts and are therefore translated more freely. Table 1 provides examples of both unwanted literal and idiomatic translations (Example 1 is from <http://www accurapid.com/goodbadfr.htm>, Examples 2 and 3 from [8]).

4. Why a Threat to Validity?

Unwanted literal translation threatens the validity of international achievement studies, because it slows down and complicates the reading process, thereby increasing the cognitive load imposed on testees: when a text is odd, unnatural, and cumbersome, it cannot be read and processed in as large chunks and as automatically and effortlessly as idiomatic and natural language (cf. [9, 10]). Rather, a considerable amount of working memory, time, and effort needs to be devoted to decoding and making meaning of the text. At the same time, less memory and energy are left for responding to the questions. This, moreover, can be expected to be the more the case, the more the text deviates from what is normal, natural, and accepted in the language. Sometimes (e.g., Example 1 in Table 1), the oddness, unnaturalness, and cumbersomeness may even result in the text being incomprehensible or misunderstood. Also, because of the

awkwardness of and the extra reading difficulty caused by text, the testee may not be as motivated to read it and to perform the tasks accompanying it.

5. Why So Difficult to Avoid?

Since unwanted literal translation endangers equivalence and validity, it is clear that great care should be taken to ensure that no test version contains such renderings. However, ensuring this is not at all easy. There are basically two reasons for this: first, literal translation is an inherent part of the translation process and the default translation strategy, because of which unwanted literal renderings are extremely common and difficult to avoid; second, there are no objective ways for identifying and assessing unwanted literal renderings.

5.1. Literal Translation: An Inherent Part of the Translation Process. Translation is a complex cognitive problem-solving and decision-making process [11, page 17]. The process consists of three main phases, during which the translator is constantly faced with translation problems which s/he needs to solve: (1) comprehension of the source text, during which the translator finds the meanings of the source text expressions; (2) transfer of the meaning of the source text into the target language, during which the translator finds target text equivalent to the source text expressions; (3) production of the target text, during which the translator decides how—for example, how literally or freely—to express source text elements in the target language [11, page 17].

However, the phases do not occur linearly, so that the translator would first comprehend the source text and only after that start producing (and making decisions on) the target text. Instead, the phases are constantly intermingling with each other so that the translator perpetually goes back and forth between the source and target text [12, pages 173–5] [13, pages 33–4]. This constant switching between the two languages, in turn, is surmised to give birth to an interlanguage or translanguage or third code, a language variant of its own which is somewhere between the two languages, sharing features of both [14, page 168] [15, pages 223–5] [16] and which is therefore inferior to them [17, pages 36–9]. A manifestation of this interlanguage is literal translation.

In addition, the presence of the source text means that when producing and making decisions on the target text, the renderings that literally imitate the source text are salient, or prominent [18], in the translator's mind. Literal translations are thus those that first come to the translator's mind, suggesting themselves as immediate, ready-made equivalents to the source text renderings [19, page 146] [20, pages 194–5] [21, page 16] [22] [23, pages 224–5]. They are what translators use—mentally and often also literally—as the first step in the translation process, the first tentative solutions to translation problems which may then be revised to something less literal. Often, however, the literal renderings are left as such, because as postulated by Levý [24, page 1179], in actual translation work (with e.g., its time limits) the translator typically has to resort to the minimax strategy and use phrasings that “promise a maximum of effect with a minimum

of effort”. Literal translation is thus the default translation strategy. It is an inherent part of the translation process. [19, page 146] [25, 26] It is a “law” of translation [23, page 275]. Literal translations (whether “wanted” or intentional, or unwanted) are thus extremely common and difficult to avoid, and extra effort is needed for translators not to translate literally and/or to get rid of unwanted literal renderings.

5.2. Lack of Objective Methods for Identifying and Assessing Unwanted Literal Translations. Another factor that makes unwanted literal translations difficult to avoid is that there are no systematic, consistent, and objective ways of identifying such renderings and assessing their impact on testees. Even though rigorous translation and quality control procedures have been developed in international achievement studies, these have been much better equipped for ensuring that the items are equivalent in meaning than ensuring that not only the items but also the stimulus texts are equally easy to understand and that they do not contain unwanted literal translations. This has most strikingly been the case with the numerous psychometric methods (statistical item analyses) used in these studies.

The above may be due to a historical fact: when the procedures were developed, they were mainly needed, for example, for psychological and social surveys, where equivalence in difficulty is typically not a concern and where the instruments usually only contain relatively short question items. Also, the notion of unwanted literal translation is extremely vague and fuzzy: it varies in kind and degree, in that it sometimes results in a text being only slightly odd and sometimes in complete nonsensicality (see Table 1); it varies across languages (and readers), and it is typically not restricted to any individual item but spreads over and taints larger portions of text (including the stimulus text). Thus, there simply are no universal criteria and methods for identifying and assessing unwanted literal translations.

Therefore, when seeking to ensure that translated tests do not contain unwanted literal translations but are in idiomatic language, one has to rely more or less exclusively on judgmental methods, (which, however, typically lack the rigor of psychometric methods). The most important part of these methods are rigorous translation procedures and practices (other methods including, e.g., cognitive laboratories and interviews with testees). The translation procedures and practices (and factors related to them, such as translators and time) thus play an extremely important role in deciding to what extent the translations contain unwanted literal translations. However, both research and experience suggest that there have been deficiencies in these procedures and that the translations have therefore often contained unwanted literal renderings [27, 28] [29, page 64] [30–33].

6. Translation Procedures in International Achievement Studies

The following provides a brief general overview of the translation procedures and practices followed in international achievement studies. However, these have differed

considerably not only between the studies but also between participating countries and over time. Therefore, for more information on the procedures and practices in each study and country, the reader is advised to refer, for example, to the websites and technical and country reports of the respective studies (although for the procedures in studies other than those conducted by the OECD and IEA, very little data are available).

6.1. Forward and Back Translation. In studies involving mainly developed countries (e.g., studies conducted by the IEA and OECD), the translations have been produced by following the forward translation procedure. However, the procedure has differed not only between the studies but also over time. For example, in the most recent IEA studies, the recommended procedure has been as follows: (1) one translator produces one target version on the basis of one (typically English) source version; (2) the target version is reviewed by a reviewer. In OECD studies, the procedure has been the following: (1) two translators produce two independent target versions—either both from English (the Programme for the International Assessment of Adult Competencies, PIAAC) or one from English and the other from French (the Programme for International Student Assessment, PISA); (2) the two target versions are merged into one national version by a translator called reconciler, who also checks that the resulting version is correct and natural. In both studies, the reviewed or reconciled version has then been verified by an international verifier.

However, not all countries have followed the recommended procedures but have slightly modified them. For example, when translating PISA materials, Finland has usually (with the exception of PISA 2006, when the procedure was as recommended) only made one translation (from English) which has then been reworked and revised by one (in PISA 2000) or two successive revisers (in PISA 2003 and 2009). Moreover, in PISA 2000 the revisions were made almost exclusively against the English source versions but in PISA 2003 and 2009 also against the French versions.

In addition to the forward translation procedure, however, in some studies also the back translation approach has been used. In this approach, the test is first translated into the target language and then back into the source language. After this, the two source language texts are compared to each other, and the quality of the target text is judged on the basis of how comparable the two source language texts are. In more recent years, the back translation approach has mainly been used in studies where the participants have come from less developed countries (e.g., SACMEQ; the Latin American Laboratory for Assessment of the Quality of Education) and where the languages may have been more exotic and therefore not known to the test developers.

In practice, the translation, review, reconciliation, and verification have mainly taken place on screen, by, for example, overwriting source language text with target language text (e.g., [34]).

6.2. Selecting Translators. The requirements for the translators have varied somewhat according to their tasks (e.g.,

translation, review, reconciliation, verification) and between the studies. However, usually countries have been advised to hire translators (hereafter used in this paper as a generic term for all those involved in translating the tests, unless otherwise specified) with a perfect command of the target language, an excellent command of the source language experience in the target culture and with students in the target population, knowledge of the subject matter, and familiarity with test development.

6.3. Translation Guidelines and Translator Training. To familiarize translators with the translation task and to help them to produce equivalent translations, international achievement studies have usually provided translators with translation and adaptation guidelines. However, the guidelines have differed between the studies in that in some studies (e.g., those conducted by the IEA and SACMEQ) they have been relatively general, with only a very few specific translation instructions. In other studies, again (e.g., those conducted by the OECD), they have contained a great number of detailed examples of the most common linguistic translation problems encountered when translating tests and advice on how to avoid them. Instructions have been given, for example, on the layout of the translations, on how to maintain the difficulty level of the vocabulary and the syntax of the text unchanged, and on how to translate the question items [34, 35]. In some studies (e.g., those conducted by the OECD) but not in all (e.g., studies conducted by the IEA), countries have also been encouraged to offer training to their translators, based on the translation guidelines. Verifiers have been trained in the International Centre.

7. Factors Having an Impact on How Literally Texts Are Translated

In Translation Studies, several factors have been found to have an impact on how literally translators translate. Among the most significant of these are the following: the purpose of the translation task and the translation guidelines, qualifications of the translators, the amount of time available, and the amount and quality of revision and the use of parallel, or comparable, texts. These are discussed in more depth in the following. For each factor, the paper first lays out what Translation Studies has to say about its impact on translation, then discusses how international achievement tests have fared with respect to it, and, finally, suggests what improvements can be made in it in order to ensure idiomatic translations.

8. Purpose of the Translation Task and Translation Guidelines

The first two factors that have an impact on how the translator translates are the purpose of the translation task and the written guidelines, or instructions, by means of which translators are typically informed about the purpose [13, 36]. The purpose is the function (or goal) of the translation (task), which governs the entire translation work

and determines how the text is to be translated [36, 37]. For example, when translating a fairy tale, the purpose is usually to produce a fictional text for children, which, in turn, requires that emphasis be put on naturalness and ease of reading. The guidelines, for their part, provide information, not only on the purpose of the translation task but often also on how (e.g., how literally or freely) the translator is to translate to attain that purpose [37, 38].

8.1. Purpose of the Translation Task. The purpose of the translation may foster literal rendering basically in two ways. First, making a literal translation may be the purpose of the translation task, as, for example, when the translator is to provide a word-for-word rendering for an exotic linguistic expression. In cases such as these, the purpose and the need to translate literally are usually also phrased clearly and unequivocally in the guidelines. Literal translations such as these are intentional (or wanted) and do not need to be avoided.

Second, the purpose may be vague, elusive, and difficult to grasp, or it may be new and strange to the translator or the guidelines may not state clearly and unequivocally how literally or freely the translator is to translate, or they may be missing altogether. When this is the case, the translator is left uncertain as to how to translate. Uncertainty, in turn, easily tempts translators into “playing safe,” avoiding risk and choosing lower-risk options. Idiomatic and “free” translations may not seem a safe choice, because they are vague; besides, they always risk being too free. Therefore, uncertainty easily results in literal translations [39, page 324] [40, page 42] [41]. Literal translations resulting from uncertainty are often unwanted and need to be avoided.

Consequently, for it to be possible to avoid unwanted literal translations in international achievement tests, it is important, first, to see to it that the purpose of the translation is clear and straightforward and that it is familiar to the translators and second, to provide translators with translation guidelines and to see to it that the guidelines state clearly and unequivocally that the translator is to avoid unwanted literal translation and to aim for idiomatic renderings. However, especially the requirement for a clear and straightforward purpose is extremely difficult, if not impossible, to accomplish. This is because the purpose when translating international achievement tests is to make all translations and different-language versions equally difficult to answer, and, yet, especially at the time of translating, there are no ways of accurately assessing the difficulty of the source and target versions. The translator therefore has no way of knowing—being certain—how difficult the source and target texts really are. Translating international achievement tests thus seems to be a task where uncertainty is necessarily present and where, accordingly, the risk of unwanted literal translations is especially high. The uncertainty has often been further aggravated by the fact that translators have not been familiar with the purpose [28, 30, 31]: Since equivalence in difficulty is typically not a purpose in any other type of translation, most translators are not trained for and used to pursuing it.

8.2. Translation Guidelines. Ways are thus needed to make the purpose more tangible, easier to grasp, and more familiar to translators. Since not much can be done to the purpose, extremely heavy demands are put on the translation guidelines. However, it seems that there have been problems in these, too. Even though most studies have provided translators with translation guidelines (or specific translator training) of some kind, in some studies (e.g., SACMEQ [42]) they have not said that translators are to aim for natural and idiomatic language and to avoid unduly literal translation. Moreover, in some studies, this need does not appear to have been stated as clearly and unequivocally as it should have been. This, in turn, appears to have been because in some studies (e.g., those conducted by the OECD) a great number of detailed linguistic instructions have been included in the guidelines so as to help translators to judge the difficulty of the texts and items and to make the target versions equivalent in difficulty to the source versions. The instructions, however, have largely consisted of directions on how to remain lexically and syntactically (and thus literally) as close to the source version as possible [34, page 15]. In practice, the guidelines thus appear to necessitate literal translation. Therefore, rather than clearly and unequivocally encouraging idiomatic translation, the guidelines seem to provide controversial messages as to how literally or freely to translate. Because of the numerous linguistic instructions—compared to the few brief recommendations for idiomatic translation—the emphasis rather seems to have been on literal translation. Such emphasis, in turn, easily lures translators into translating literally.

The great number of specific linguistic instructions and the strong emphasis in the OECD guidelines on literal resemblance can be expected to be a problem in translations into non-Indo-European languages, in particular. This is because the instructions are mainly written from the point of view of Indo-European languages (and, specifically, English and French) and their syntactic and vocabulary structures. Therefore, when translating into these languages, following the instructions and staying close to the source version often work quite well, yielding completely natural literal translations. However, when translating into non-Indo-European languages (e.g., Finnish), where the syntactic and vocabulary structures differ from those of English and French, the attempt to follow the instructions easily leads to unwanted literal translations [28] [43, page 28].

Thus, adjustments may be needed in the translation guidelines. The first step, however, is to see to it that translators in all studies receive translation guidelines. The next step is to ensure that the guidelines say clearly that the main concern is to translate idiomatically and that this is because unnaturalness inevitably leads to nonequivalence in difficulty. At the same time, explicit warnings—supported by illustrative examples—are needed against unwanted literal translation. This is to counteract both the strong effect the specific word—and sentence—level instructions have in the opposite direction and the tendency of translators to translate literally. Some specific linguistic translation instructions do seem to be needed to make the elusive, unique and strange purpose of the translation task clearer

and to render it easier for translators to assess the difficulty of the items and texts. However, could the number of the instructions perhaps be reduced, as suggested, for example, by, Dept et al. [44, page 165]? More research is needed on this. Also, when providing such instructions, it is important to remind translators that because of differences between languages, the instructions do not always apply and that, therefore, slavishly following them easily leads to unwanted literal translations. It may even be good to prepare separate, customized instructions for translation into languages that are very dissimilar to English and French (as has already been done in PISA for translation into Arabic and Chinese).

Another way of making the purpose of the translation of international achievement tests more tangible, easier to grasp, and more familiar to translators is to provide translators with explicit training on it. This has been the practice in some but not in all studies. In the training, the emphasis should also be on the need to translate idiomatically.

9. Qualifications of the Translators

Translators and their qualifications have a great impact on how they translate. For example, translators with little or no experience and training in translation and little or no knowledge of the theory and principles of translation often have a naïve, distorted view of translation. Unlike their more qualified peers, who have been trained to take into account target text readers, translate meanings (rather than e.g., words), and aim for idiomaticity and naturalness, less qualified translators often erroneously see translation as a formal, word-for-word transfer, where they are expected to follow closely the source text and translate literally [11, page 31] [12, pages 171-2] [20, page 199] [21, page 12] [23] [45, page 166] [46, pages 113-4] [47, page 221] [48]. Less experienced and less qualified translators also seem more uncertain than qualified translators [11] and may therefore be tempted to “play safe” and translate literally [11, page 36] [39, page 324] [40, page 42].

Consequently, a first step in ensuring that translations in international achievement tests do not contain unwanted literal renderings is to see to it that only qualified translators are used to translate, revise and verify them and that they have a good knowledge of the source language or languages (reconcilers and verifiers in PISA) and especially of the target language, are well versed in the subject matter, and are familiar with translation theory and the general principles of translation. However, it seems that this has not always been the case, but, rather, that the translators have often lacked important qualifications [28, 30, 31, 49].

More attention thus needs to be paid to ensuring that the translators really are qualified. A practical way of doing this is to test them (see also, e.g., [5] [50, pages 12-5]). The qualifications should also be clearly mentioned in the requirements for the translators. What also helps translators to avoid unwanted literal translations is to make it possible for them to work in teams and discuss with subject matter experts, for example, (see also, e.g., [5, 50]). However, such discussions take a lot of time, and therefore it is necessary to reserve time for them in the testing and translation schedule.

10. The Amount of Time Available

Time likewise plays a role in how literally the translator translates. Since literal renderings are the first that come to the translator’s mind and the first that are used as translation equivalents, extra effort and time are often needed so that the translator—or reviser—can get rid of the literal renderings, elaborate on the text, and make it more natural. However, when in a hurry or under time pressure, the translator lacks cognitive resources (cf. [51]) and has no time for problem solving [45]. S/he has no time to be creative (see, e.g., [52, page 444]) and to invent more idiomatic expressions [19, 53]; also, s/he has no time to do research and consult others [54]. Instead, s/he has to resort to the minimax strategy and be satisfied with the solutions that first and most effortlessly come to his or her mind (literal translations), even though these may not be the best solutions [40, page 43] [45] [55, 56, page 507].

Thus, to avoid unwanted literal translations in international achievement studies, it is important to see to it that (the) translators (translating the tests) have sufficient time to do their job. However, findings suggest that this has not always been the case but that, rather, the translators have often had to work under very tight timelines and time pressure [28, 30, 31, 57, 58].

There are, roughly speaking, two ways in which it can be ensured that translators have sufficient time to translate the tests. The first is to hire enough translators for each translation phase. However, finding several qualified translators may not always be easy (B. Halleux-Monseur, personal communication, January 24, 2008). The second way, then, is to allot more time to translation in the testing schedule. However, this requires that major changes be made in the testing cycles.

11. The Amount and Quality of Revision and the Use of Parallel Texts

How much a translation contains literal rendering is also dependent on how it is revised or checked for correction and improvement [55, 56, page 17]. In international achievement studies, the way the translations are revised basically depends on which of the two major translation approaches is followed: forward or back translation. In forward translation, there are, in principle, two phases during which the translations can be revised: review (e.g., IEA studies) or reconciliation (OECD studies), during which the translations are checked by a national reviewer or reconciler and verification, during which they are revised by an international verifier. For the sake of brevity, this paper only discusses national revision, even though the same principles, of course, also apply to verification (for more information on verification, see, e.g., [44]). Since the review or reconciliation phase is typically not reserved for revision alone, but also includes another task, that of merging together two parallel target versions, these two tasks are here discussed together. In the other major translation approach, back translation, the revision largely consists in comparing the back translated versions to the source versions.

11.1. Revision. Revising can have a great effect on whether or to what extent a translation contains unwanted literal renderings. How a translation is revised, in turn, depends on at least five largely interrelated factors, most of which also have an impact on how texts are actually translated: the purpose of and guidelines specifying the translation task, the method of revising, the time spent on revising, the person(s) revising, and the medium for revising (paper or screen). Deficiencies in any of these easily result in the reviser not being able to spot and correct unwanted literal translations.

As in the case of translation proper, the purpose of the translation task and the translation guidelines have a central role in determining how a text is revised. If the purpose is vague or difficult to grasp or if the reviser is not familiar with it and/or if the guidelines do not say clearly and unequivocally that the translation is meant to be idiomatic, the reviser is left uncertain as to how to proceed and may therefore be tempted to “play safe” and accept literal translations (cf. [39, page 324]; see also [40, page 42] [41]).

Obviously, the method of revision plays a huge part in how a translation is revised. Properly revising a translation requires that the translation be checked, for example, for (semantic) faithfulness to the source text, grammatical correctness, naturalness and idiomaticity, and correctness of style. However, these cannot all be checked at the same time but necessitate several separate revisions. For example, there should be a separate revision for checking the translation for faithfulness to the source version and another for checking it for idiomaticity. If both were checked at the same time, the presence of the source text and its wordings would be strongly salient in the reviser’s mind, which, in turn, would make him or her blind to unwanted literal renderings. Another purely monolingual revision is therefore needed to check translations for idiomaticity. (cf. [19, pages 32, 233] [55, 56, page 147]) By the same token, the monolingual revision should preferably precede the bilingual revision (if both are made by the same person) [8, 57, 58].

The previously mentioned partly explains why, for example, back translation is not effective when checking translations for unwanted literal renderings: when using back translations, the reviser mainly concentrates on the back translated versions and their semantic faithfulness to the source texts, with much less attention paid to the translation and whether it is in idiomatic language or contains unwanted literal renderings [8]. Another factor that makes back translation problematic is that it may affect, not only the reviser but also the translator making the translation draft even encouraging him or her to translate literally: When the translator knows that the translation will be translated back into the source language and that it will be assessed, not on the basis of the translation and its idiomaticity and naturalness, but on the basis of the back translated text and its correspondence with the source text, s/he may think—in harmony with the minimax strategy—that there is no point in pursuing idiomaticity, because it would only make the task more difficult both for him or her and for the back translator [59].

Making several separate revisions, of course, takes a lot of time. Therefore, a lack of time easily results in the reviser

making only one revision during which s/he tries to check everything: both the faithfulness of the translation to the source text and its naturalness, for example. However, trying to concentrate on both the source and target text makes it difficult for the reviser to spot and correct unwanted literal translations [19, pages 32, 233] [55, 56, page 147]. Or because of the lack of time, s/he may be tempted to accept the literal renderings, because they are the quickest and easiest choice (the minimax strategy).

Understandably, the revisers and their qualifications also have an impact on how the revision is made. For example, revisers with a deficient knowledge of the target language will not be able to ensure that the translations are in natural and idiomatic language. Or revisers with no academic translator and/or reviser training may not know how to revise and what is involved in revision; also, they may think that literal translations are good translations. The number of revisers likewise affects revision. For example, when there is only one reviser, s/he has to check everything. However, when there are several revisers, they can divide the tasks so that, for instance, one of them only concentrates on the monolingual revision. Having several revisers also means that there are more eyes to spot unwanted literal translations. [55, 56]

Still another factor that has an effect on revision and to what extent revisers are able to spot and do away with unwanted literal renderings is whether the revision is made on paper or screen: spotting unwanted literal renderings is more difficult on screen [19, page 144] [25, page 36] [55, 56, page 101].

11.2. Parallel Texts. Parallel texts can help the translator to avoid unwanted literal renderings. They can help the translator to get a fuller understanding of the meaning of the source text, which, in turn, is a prerequisite for him or her to be able to make a natural and fluent translation. In contrast, deficient comprehension (and the ensuing uncertainty) easily leads to risk aversion and to literal and often incomprehensible translations [39, page 324] [40, page 42]. Also, parallel texts help the translator to see that typically there is not just one but several ways in which an idea can be expressed and that these can differ enormously and, yet, all be correct. This can encourage the translator not to use literal renderings but to choose more idiomatic expressions.

However, the use of parallel texts may also foster unwanted literal translation. This is the case when the texts need to be merged together. Merging texts together is a complex cognitive process, which involves, among other things, comparing the texts to each other, taking out ideas and extracts from them and putting these together. However, when ideas and extracts from different sources are put together, it cannot be assumed that the resulting text would automatically be correct, coherent, harmonious, and in good language. Rather, as a last step in the merging process, the resulting text also needs to be carefully revised and finalized. This need, moreover, is understandably the greater, the more different the parallel texts are from each other (as, e.g., when they are translated from different languages).

Thus, the whole process of merging, if done properly, is a complex cognitive process and requires a lot of time.

Therefore, if the translator is in a hurry, s/he often has to compromise on the quality of the revision. Also, when time is scarce, the source text and its wordings (to which the person doing the merging needs to compare the target text) will necessarily remain in his or her mind when s/he starts to check the translation for naturalness and make him or her blind to unwanted literal renderings.

11.3. Revision and Parallel Versions in International Achievement Studies. In the context of international achievement studies, the most important lessons from the previously mentioned are the following. First, if we want to avoid unwanted literal translations in these studies, it is necessary to have all translations properly revised and finalized. This, in turn, requires the following: that the purpose or goal of translating the tests is made clear to reconcilers and reviewers by providing them with clear translation guidelines and translator training; that the translation approach makes it possible for revisers to focus on the translations (not on, e.g., back translations) and on making them idiomatic; that reconcilers and reviewers have sufficient time to make the revision (several revision rounds); that they are qualified. In addition, it is beneficial to use several successive revisers and make the revision on paper. Second, it may be good to make parallel target versions—provided the other tasks involved in using them (e.g., comparing the versions to each other and merging them together) do not complicate the revising and finalizing of the translations, by, for example, consuming so much time that the reviser does not have sufficient time to make the revision.

11.3.1. Revision. However, it seems that in international achievement studies sufficient attention has not always been paid to the revision and finalizing of the translations (e.g., [60]; see also [61]) and that therefore it may have failed to spot and correct unwanted literal translations. Even though no research proper exists on why this may have been so, there are findings [28] suggesting that the reasons have been largely the same as those when actually translating a text: that the purpose of the translation task has been vague and strange and that the guidelines have not always been quite clear and unequivocal; that the method of revision has not always been efficient; that the revisers have not always had sufficient time to make the revision; that the revisers have not always been qualified. In addition to this, the deficiencies may also have been due to the fact that in international achievement studies the revisions have mainly been done on screen.

Consequently, to avoid unwanted literal translations in international achievement tests, it is imperative that more attention be paid to the revising and finalizing of the translations. This involves, for example, the following. First, making sure that the revisers understand the purpose and specifics of the translation task, by providing them with written translation guidelines and translator training which say clearly that the goal is to make translations that are natural and in idiomatic target language. Partly customized guidelines (and training) may also be needed for translation into more remote languages. Second, using a translation approach which makes it possible for revisers to concentrate

on the target text—this typically rules out back translation, unless it is accompanied by a separate revision for idiomatic language [59, page 39]—and reminding them of the need to make several revision rounds, of which one should focus on ensuring the naturalness and idiomaticity of the translations. Third, allotting so much time to revision and using so many parallel revisers that the revisers have sufficient time to make several separate revision rounds, to be creative and to seek for idiomatic expressions, to revise on paper, and to discuss with subject matter experts, when necessary. Fourth, making sure—by means of a test, for instance—that the revisers have an excellent command of the target language and that they are well versed in the principles of translation. Fifth, strongly encouraging revisers to make the revision at least partly on paper. This has become increasingly important today, when more and more of the translation work is done in electronic environments.

In addition to all this, however, as part of the revision process, it would be good to have pilot testees (and other outsiders) read the translations (with fresh eyes, with no negative influence from the source versions), complete the test, and comment on the language in it. Cognitive laboratories such as these are the only way to find out whether or to what extent unwanted literal renderings really affect testees and their performance (cf. [62]).

11.3.2. Parallel Versions. When translating international achievement tests, two types of parallel target versions have been used: those that are translated from one and those that are translated from two source versions. For the sake of brevity, the following mainly limits itself to the latter. This is because when using two different-language source versions, the problems may be expected to be greater. Also, more findings are available on this practice. However, the principles also largely apply to cases where the target versions are rendered from only one source version.

International achievement studies have differed considerably in whether or not they have used two source versions. For example, in IEA studies, it was previously recommended that each country makes two target versions from one (or two) source version(s); today, the recommendation is to make only one target version. In the International Adult Literacy Survey (IALS), too, some countries (e.g., Finland) made their translations on the basis of two source version. However, today, parallel source versions are only used in OECD PISA studies (in, e.g., the OECD PIAAC, parallel target versions are made from one source version).

The experiences as to whether the use of two source versions has helped to avoid unwanted literal renderings or not have also varied. In some cases, the experiences have been mainly positive. For example, in the PISA 2000 field trial, verifiers reported that those test versions that had been translated from two source versions contained fewer unwanted literal translations than those that had been rendered from only one version [29]. Also, for instance, Denmark has found the procedure to be beneficial (J. Mejding, personal communication, September 16, 2011).

However, there are also experiences which suggest that the use of two source versions may foster unwanted literal

translation. For example, when translating the IALS materials into Finnish (by means of double-translation from two languages), the reconciler felt that the use of the two source versions was complicated. The two Finnish target versions, one of them based on the English and the other on the French source version, were often so different that much more time would have been needed to make the resulting versions coherent, fluent, and idiomatic. Therefore, Finland decided not to follow the procedure in PISA but to modify it slightly so that more emphasis would be put on the revising and finalizing of the target versions. (P. Linnakylä, personal communication, November 14, 2008). In PISA 2000, Finland thus only made one translation from each (English) source version, which was then revised by another national translator. In PISA 2003, the procedure was developed further so that each draft was revised by two successive revisers; the first of whom also cross-checked the drafts against the French source versions.

This, as commented by the Finnish translators translating the PISA 2009 materials [28], has had both its cons and pros. The main disadvantage has been an increase in time pressure. When the total amount of time reserved for translation has remained the same and when, at the same time, this time has had to be divided between three (instead of two) successive national translators and revisers, each translator and reviser has had less time to do his or her job. In practice, however, the time pressure seems to have mainly centered on the first reconciler, who has had several tasks to accomplish. By contrast, the second reconciler, for example, who has been able to concentrate more or less exclusively on revising and finalizing the Finnish versions, does not appear to have had suffered from time pressure. Mainly, then, the procedure has been found to be beneficial. The use of the two different-language source versions has helped translators to find different ways of expressing the same thing and the correct meanings of words with several meanings, which, in turn, has helped them to avoid unwanted literal renderings; also, the procedure has made it possible to pay proper attention to the revising and finalizing of the Finnish versions, without interference from the source languages. Interestingly, in international verification, the linguistic quality of Finnish PISA tests has also been judged to be very high.

Experiences from Sweden also suggest that the double translation and reconciliation procedure may make it difficult for the reconciler to properly revise and finalize the translations. For example, reports from verifiers checking Sweden's and Finland's Swedish PISA materials (Swedish being one of the two languages in which the tests have been provided in Finland) suggest that the linguistic quality of Sweden's materials has usually been clearly lower than that of Finland's Swedish materials: they have contained, among other things, more errors and less fluent and natural language. However, basically the materials have been the same, because Finland has borrowed materials from Sweden and only adapted them for use in Finland. In practice, then, the only difference has been that in Finland more time has been spent on making the revision. During this extra time, the materials have undergone an extra revision round, during which they have been checked by an extra person, who has

been able to concentrate solely on the final Swedish versions (without interference from the source versions or the first Swedish drafts) and on finalizing them. Thus, it seems that in at least some countries, the use of the two source versions has rendered it difficult for reconcilers to find sufficient time to make a proper revision and to ensure that the translations do not contain unwanted literal translations.

Therefore, if parallel source versions are used, it is important to see to it that reconcilers have sufficient time also to properly revise and finalize the resulting target versions. In practice, this might be done, for example, so that each country hires so many reconcilers that each reconciler only has a small number of texts to revise. However, the problem with this solution is that since the requirements for reconcilers are so high and unique, it may not be easy to find several persons who would meet the requirements. Another slightly better solution, then, would be to allot more time in the translation schedule to the reconciliation phase. However, this solution also has its problems. Since the reconciler first needs to examine the two target and two source versions and merge the two target versions into one, it is more than likely that all the different versions will continue to have an impact (interference) on him or her also while revising and blind him or her to unwanted literal renderings. Therefore, the best option would be to split the reconciliation phase into two so that the first phase would consist in merging together the two target versions, whereas the second phase would be devoted to revising and finalizing the resulting target versions. However, here the problem is that splitting the reconciliation phase into two and making revision a phase of its own require major changes in the timing of the testing procedures.

All in all, however, given the contradictory experiences gained in using two different-language source versions, it is obvious that more research is needed on the procedure and on how it affects the revising, finalizing, and idiomaticity of translated test versions. For example, it is important to know whether there have been differences in how the procedure has actually been implemented in the participating countries (e.g., whether some countries have used more or less exclusively only one of the target versions, with only a few small extracts taken from the other target version and whether other countries have used the two target versions more equally) and which of these would be the best able to guarantee idiomatic translations. Research is likewise needed to find out whether or to what extent the use of two target versions—which, like the use of two source versions, also requires merging together two parallel versions—has a similar negative impact on revision and idiomaticity as the use of two source versions or whether this impact is somewhat weaker (because when the target versions are translated from one and the same source version, they may be expected not to be as different from each other as when they are rendered from two source versions, and therefore also merging them into one coherent and idiomatic whole may be expected to be easier and less time consuming; also, when the reconciler only has one source version to which to compare the translations, s/he not only has slightly more time to revise the translations but may also be expected to be

less affected by interference). Finally, research is needed on all the other procedures actually followed in the participating countries (e.g., the procedure followed in Finland) to find out how to what degree they have been successful in producing idiomatic translations and why. In practice, this might be done, for example, by asking translators, revisers, and verifiers working in or for the various countries about their experiences in following the procedures and by making comparisons between translations produced by following the various translation procedures.

12. Conclusion

This paper discussed unwanted literal translation in international achievement studies. The purpose was twofold: to increase awareness of the threat unwanted literal translation poses to the validity of these studies and to discuss problems there have been when translating these tests which may have endangered their idiomaticity and to provide suggestions on how to improve the translation work so as to ensure as idiomatic translations as possible.

The paper showed that unwanted literal translation threatens the validity of international achievement studies, because in these studies the instruments need to be equivalent, not only in meaning but also in difficulty and because unwanted literal translations make attaining this goal impossible. By making texts odd and unnatural, they complicate and slow down the reading and response process and decrease the motivation of the testee to read the text and to answer the questions, thereby putting testees at an unequal position. What further aggravates the problem is that literal translation (whether wanted and unwanted) is the default translation strategy and as such extremely common and difficult to avoid. Also, there are no psychometric methods for systematically and objectively identifying unwanted literal renderings in these tests and assessing their effects on testees. Thus, to avoid unwanted literal renderings, the studies have to rely more or less exclusively on judgmental methods and, more specifically, on rigorous translation procedures and practices.

However, the paper showed that there have been problems in these procedures and practices which have made it difficult to attain idiomaticity. For example, in these studies the purpose of the translation task is necessarily vague and often also strange to the translators and revisers. In addition, the translation guidelines have not always made it unequivocally clear that the goal has been to produce idiomatic translations. As a result, translators and revisers have often been left uncertain as to how to translate, which, in turn, easily results in risk aversion and literal translation. In addition, translators and revisers have often lacked qualifications or been inexperienced, because of which they may have had a naïve and false conception of translation (simple word-to-word mappings) or they may have been uncertain and therefore resorted to literal renderings. Also, they have often had to work in a hurry, because of which they may have had to accept the literal renderings which have first come to their minds. Finally, when revising the translations, revisers have not always been able to pay

sufficient attention to the naturalness and idiomaticity of the translations, because they have had to use the back translation method, because they have had to merge together two parallel versions or because they have worked on screen. What has made avoiding unwanted literal translations even more difficult is that often there have been problems not only in one but several factors at the same time (e.g., incompetent translators, translating in a hurry, and instructed to keep the syntactic structures unchanged).

To help increase the idiomaticity and equivalence of different-language versions of international achievement tests, the paper suggested the following.

- (i) Purpose: Make the purpose of the translation task as clear and familiar to the translators and revisers as possible, by providing them with translation guidelines and hands-on training.
- (ii) Translation Guidelines and Translator Training: Say clearly in the guidelines and training that the main concern is to translate idiomatically and explain why this is so (e.g., by means of examples). Prepare customized instructions for translation into more remote languages.
- (iii) Translators and Revisers: Make sure (e.g., by means of a test) that the translators and revisers are qualified, with a good knowledge of the source and especially the target language, the subject matter and translation theory. Make it possible for translators and revisers to discuss with subject matter experts, for example.
- (iv) Time: Allot sufficient time in the testing schedule to translation, team discussions, and revision.
- (v) Revision and Finalizing: Use a translation approach which allows revisers to focus on the target text and on making it idiomatic (this often rules out back translation). If two parallel versions are used, make the revising and finalizing of the translations a phase of its own. Advise revisers to make several revision rounds and to pay special attention to idiomatic target language. Encourage revisers to revise on paper. Conduct cognitive laboratories.

In addition to this, the paper also left open or raised new questions that need to be addressed in future research. For example, research is needed to find the ideal number and ideal way of presenting specific translation instructions. This could be done, for example, by making several versions of the translation guidelines and asking several translators (into, e.g., Arabic, Chinese, Greek, Finnish, Icelandic, Korean, Russian) with comparable qualifications to translate the same materials into their native languages by using the different versions and then comparing their experiences and the resulting translations. Research is also needed on how the use of the two target or source versions affects the revising, finalizing, and idiomaticity of the translations and how these two tasks can be best combined. This necessitates comparisons between the translation procedures followed in the various organizations and participating countries

(e.g., Finland) and, especially, between the translations made when following the procedures.

As our understanding of the significance and principles of translation grows, it helps us to improve the procedures and practices followed when translating international achievement tests and to produce more idiomatic and equivalent translations. Conversely, however, this also means that the practices followed when translating, for example, the very first tests were not as developed as they are today (e.g., in the first OECD translation guidelines, much more emphasis was put on close faithfulness to the source version) and that the translations made at that time are not as idiomatic and equivalent as they are today. This, of course, casts doubts on the validity not only of the early studies but also of all those studies where the early materials have been or will be used as anchors (to provide trend data). Naturally, nothing can be done to improve the validity of the past studies. However, by making a close linguistic examination of the early translations, we can decide whether they are of a sufficiently high quality to be used in future studies and in this way ensure the validity of the future studies.

Finally, the fact that translations often contain unwanted literal renderings and that they tend to be inferior to untranslated texts means that testees responding to translated versions are at a disadvantage, when compared to testees responding to untranslated (e.g., English) versions. This has led some researchers (e.g., [63]) to suggest that materials in international achievement studies be indigenous (the comparability of these materials would be ensured by analyzing them against a given set of criteria of text and item difficulty). This approach would undoubtedly improve the authenticity and idiomaticity of the national versions and in this way increase equivalence. At the same time, however, it also poses huge challenges to equivalence. Future studies could examine whether it would be possible to combine the strengths of these two approaches (translated versus untranslated test versions).

Acknowledgments

The preparation of this paper has been supported by the Academy of Finland (Grant no. 126855).

References

- [1] R. Hambleton and A. Zenisky, "Translating and adapting tests for cross-cultural assessment," in *Cross-Cultural Research Methods in Psychology*, D. Matsumoto and F. van de Vijver, Eds., pp. 46–73, Cambridge University Press, Cambridge, Mass, USA, 2011.
- [2] J. Sweller, "Cognitive load during problem solving: effects on learning," *Cognitive Science*, vol. 12, no. 2, pp. 257–285, 1988.
- [3] R. Rueda, "Cultural perspectives in reading: theory and research," in *Handbook of Reading Research IV*, M. Kamil, P. D. Pearson, E. B. Moje, and P. Afflerbach, Eds., pp. 84–104, Routledge, New York, NY, USA, 2011.
- [4] R. Tourangeau, L. Rips, and K. Rasinski, *The Psychology of Survey Response*, Cambridge University Press, Cambridge, Mass, USA, 2000.
- [5] J. Harkness, A. Villar, and B. Edwards, "Translation, adaptation and design," in *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*, J. Harkness, M. Braun, B. Edwards et al., Eds., pp. 117–140, John Wiley & Sons, Hoboken, NJ, USA, 2010.
- [6] E. Nida, *Toward a Science of Translating*, Brill, Leiden, The Netherlands, 1964.
- [7] E. Nida and C. Taber, *The Theory and Practice of Translation*, Brill, Leiden, The Netherlands, 1969.
- [8] M. Larson, *Meaning-Based Translation. A Guide to Cross-Language Equivalence*, University Press of America, Lanham, Md, USA, 2nd edition, 1998.
- [9] W. Kintsch and T. A. van Dijk, "Toward a model of text comprehension and production," *Psychological Review*, vol. 85, no. 5, pp. 363–394, 1978.
- [10] T. van Dijk and W. Kintsch, *Strategies of Discourse Comprehension*, Academic Press, Orlando, Fla, USA, 1983.
- [11] E. Angelone, "Uncertainty, uncertainty management and metacognitive problem solving in the translation task," in *Translation and Cognition*, G. Shreve and E. Angelone, Eds., American Translators Association Scholarly Monograph Series 15, pp. 17–40, Benjamins, Amsterdam, The Netherlands, 2010.
- [12] J. Danks and J. Griffin, "Reading and translation. A psycholinguistic perspective," in *Cognitive Processes in Translation and Interpreting*, J. Danks, G. Shreve, S. Fountain, and M. McBeath, Eds., Applied Psychology. Individual, Social, and Community Issues Vol. 3, pp. 161–175, Sage, Thousand Oaks, Calif, USA, 1997.
- [13] C. Nord, "Loyalty and fidelity in specialized translation," *Confluências—Revista de Tradução Científica e Técnica*, vol. 4, pp. 29–41, 2006.
- [14] W. Frawley, "Prolegomenon to a theory of translation," in *Translation: Literary, Linguistic, and Philosophical Perspectives*, W. Frawley, Ed., pp. 159–175, Associated University Press, London, UK, 1984.
- [15] G. Toury, "Interlanguage and its manifestation in translation," *Meta*, vol. 23, pp. 223–231, 1979.
- [16] A. Al-Hassnawi, "Translanguage vs. interlanguage: exploration in translation strategies," *Translation Journal*, vol. 14, no. 3, 2010, <http://translationjournal.net/journal//53translanguage.htm>.
- [17] A. Chesterman, "Beyond the particular," in *Translation Universals? Do They Exist*, A. Mauranen and P. Kujamäki, Eds., pp. 33–49, Benjamins, Amsterdam, The Netherlands, 2004.
- [18] R. Giora, "On the priority of salient meanings: studies of literal and figurative language," *Journal of Pragmatics*, vol. 31, no. 7, pp. 919–929, 1999.
- [19] B. Englund Dimitrova, *Expertise and Explicitation in the Translation Process*, Benjamins, Amsterdam, The Netherlands, 2005.
- [20] P. Kujamäki, "What happens to 'unique items' in learners' translations? 'Theories' and 'concepts' as a challenge for novices' views on 'good translation,'" in *Translation Universals. Do They Exist?* A. Mauranen and P. Kujamäki, Eds., pp. 187–204, Benjamins, Amsterdam, The Netherlands, 2004.
- [21] S. Tirkkonen-Condit, "Process research: state of the art and where to go next," *Across Languages and Cultures*, vol. 3, no. 1, pp. 5–19, 2002.
- [22] S. Tirkkonen-Condit, "Unique items—over- or under-represented in translated language?" in *Translation Universals. Do They Exist?* A. Mauranen and P. Kujamäki, Eds., pp. 177–184, Benjamins, Amsterdam, The Netherlands, 2004.
- [23] G. Toury, *Descriptive Translation Studies and Beyond*, Benjamins, Amsterdam, The Netherlands, 1995.

- [24] J. Levý, "Translation as a decision process," in *To Honor Roman Jakobson: Essays on the Occasion of His Seventieth Birthday II*, J. Levý, Ed., pp. 1171–1182, Mouton, The Hague, The Netherlands, 1967.
- [25] B. Englund Dimitrova, "Literal translation in the translation process of professional translators," *Gothenburg Studies in English*, vol. 90, pp. 29–39, 2005.
- [26] S. Tirkkonen-Condit, J. Mäkisalo, and S. Immonen, "The translation process—interplay between literal rendering and a search for sense," *Across Languages and Cultures*, vol. 9, no. 1, pp. 1–15, 2008.
- [27] I. Arffman, *In search of equivalence—translation problems in international literacy studies [M.S. thesis]*, University of Jyväskylä, Jyväskylä, Finland, 2002.
- [28] I. Arffman, "Translating international achievement tests: translators' view," Finnish Institute For Educational Research Reports 44, University of Jyväskylä, Jyväskylä, Finland, 2012, <http://ktl.jyu.fi/img/portal/22708/g044.pdf>.
- [29] A. Grisay, "Translation and cultural appropriateness of the test and survey material," in *PISA, 2000 Technical Report*, R. Adams and M. Wu, Eds., pp. 57–70, OECD, Paris, France, 2002.
- [30] R. Hambleton, "Adapting achievement tests into multiple languages for international assessments," in *Methodological Advances in Cross-National Surveys of Educational Achievement*, A. Porter and A. Gamoran, Eds., pp. 58–79, National Academy Press, Washington, DC, USA, 2002.
- [31] R. Hambleton, "Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures," in *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*, R. Hambleton, P. Merenda, and C. Spielberger, Eds., pp. 3–38, Erlbaum, Mahwah, NJ, USA, 2005.
- [32] OECD, *Country-by-Country Report on the Quality of the National Version(s) of the PISA Material*, OECD, Paris, France, 2001.
- [33] M. Wu, "A critical comparison of the contents of PISA and TIMSS mathematics assessments," 2009 <https://edsurveys.rti.org/PISA/documents/WuA.Critical.Comparison.of.the.Contents.of.PISA.and.TIMSS.psg.WU.06.1.pdf>.
- [34] OECD, "PISA, 2009 translation and adaptation guidelines," in *National Project Managers' Meeting*, Dubrovnik, Croatia, September 2007.
- [35] OECD, *National Project Manager's Manual*, OECD, Paris, France, 1999.
- [36] K. Reiss and H. Vermeer, *Grundlegung Einer Allgemeinen Translations Theorie*, Niemeyer, Tübingen, Germany, 1984.
- [37] H. J. Vermeer, "Skopos and commission in translational action," in *Readings in Translation Theory*, A. Chesterman, Ed., pp. 173–187, Finn Lectura, Helsinki, Finland, 1989.
- [38] B. Hatim and I. Mason, *The Translator as Communicator*, Routledge, London, UK, 1997.
- [39] A. Pym, "On Toury's law of how translators translate," in *Beyond Descriptive Translation Studies: Investigations in Homage to Gideon Toury*, A. Pym, M. Schlesinger, and D. Simeoni, Eds., pp. 311–328, Benjamins, Amsterdam, The Netherlands, 2008.
- [40] A. Chesterman, "Why study translation universals?" in *Kiasm*, R. Hartama-Heinonen and P. Kukkonen, Eds., University of Helsinki, Department of Finnish, Finno-Ugrian and Scandinavian Studies, Helsinki, Finland, 2010, *Acta Translatologica Helsingiensia*, vol. 1, pp. 38–48.
- [41] M. Jensen, *Professional translators' establishment of skopos—a "brief" study [M.S. thesis]*, Aarhus School of Business, Aarhus University, Aarhus, Denmark, 2009.
- [42] Southern Africa Consortium for Monitoring Educational Quality (SACMEQ), *Manual for National Research Co-Ordinators*, International Institute for Educational Planning, Paris, France, 1995.
- [43] A. Grisay, J. H. A. L. de Jong, E. Gebhardt, A. Berezner, and B. Halleux-Monseur, "Translation equivalence across PISA countries," *Journal of Applied Measurement*, vol. 8, no. 3, pp. 249–266, 2007.
- [44] S. Dept, A. Ferrari, and L. Wäyrynen, "Developments in translation verification procedures in three multilingual assessments: a plea for an integrated translation and adaptation monitoring tool," in *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*, J. Harkness, M. Braun, B. Edwards et al., Eds., pp. 157–173, John Wiley & Sons, Hoboken, NJ, USA, 2010.
- [45] A. Jensen, *The Effects of Time on Cognitive Processes and Strategies in Translation*, Copenhagen Business School, Faculty of Modern Languages, Copenhagen, Denmark, 2000.
- [46] A. Jensen and A. Jakobsen, "Translating under time pressure," in *Translation in Context. Selected Contributions From the EST Congress, Granada 1998*, A. Chesterman, N. Gallardo San Salvador, and Y. Gambier, Eds., pp. 105–116, Benjamins, Amsterdam, The Netherlands, 2000.
- [47] R. Jääskeläinen, "Are all professionals experts?" in *Translation and Cognition*, G. Shreve and E. Angelone, Eds., American Translators Association Scholarly Monograph Series, 15, pp. 213–227, Benjamins, Amsterdam, The Netherlands, 2010.
- [48] R. Kim, "Use of extralinguistic knowledge in translation," *Meta*, vol. 51, no. 2, pp. 284–303, 2006.
- [49] I. Karg, *Mythos PISA. Vermeintliche Vergleichbarkeit und die Wirklichkeit eines Vergleichs*, V&R Unipress, Göttingen, Germany, 2005.
- [50] European Social Survey, *ESS Round 5 Translation Guidelines*, European Social Survey GESIS, Mannheim, Germany, 2010.
- [51] D. Zakay, "The impact of time perception processes on decision-making under time stress," in *Time Pressure and Stress in Human Judgment and Decision Making*, O. Svenson and A. J. Maule, Eds., pp. 59–72, Plenum, New York, NY, USA, 1993.
- [52] M. Fontanet, "Time of creativity in translation," *Meta*, vol. 50, no. 2, pp. 432–447, 2005.
- [53] R. Jääskeläinen, "Hard work will bear fruit? A comparison of two think-aloud protocol studies," *Meta*, vol. 16, pp. 60–74, 1996.
- [54] R. Mackenzie, "Creative problem-solving and translator training," in *Translator's Strategies and Creativity*, A. Beylard-Ozeroff, J. Králová, and B. Moser-Mercer, Eds., pp. 201–206, Benjamins, Amsterdam, 1998.
- [55] H. Krings, *Was in den Köpfen von Übersetzern vorgeht. Eine empirische Untersuchung der Struktur des Übersetzungsprozesses an fortgeschrittenen Französischlernern*, Narr, Tübingen, Germany, 1986.
- [56] B. Mossop, *Revising and Editing for Translators*, St. Jerome, Manchester, UK, 2007.
- [57] R. Hambleton and G. Berberoglu, *TIMSS Instrument Adaptation Process: A Formative Evaluation*, University of Massachusetts, School of Education, Amherst, Mass, USA, 1997.
- [58] A. Künzli, "Translation revision: a study of the performance of ten professional translators revising a legal text," in *Doubts and Directions in Translation Studies*, Y. Gambier, M. Schlesinger, and R. Stolze, Eds., pp. 115–126, Benjamins, Amsterdam, The Netherlands, 2007.

- [59] F. van de Vijver and K. Leung, *Methods and Data Analysis for Cross-Cultural Research*, Sage, Thousand Oaks, Calif, USA, 1997.
- [60] G. Solano-Flores, L. Contreras-Niño, and E. Backhoff-Escudero, "Translation and adaptation of tests: lessons learned and recommendations for countries participating in TIMSS, PISA and other international comparisons," *Revista Electrónica de Investigación Educativa*, vol. 8, no. 2, 2006, <http://redie.uabc.mx/vol8no2/contents-solano2.html>.
- [61] G. Solano-Flores, E. Backhoff, and L. Contreras-Niño, "Theory of test translation error," *International Journal of Testing*, vol. 9, pp. 78–91, 2009.
- [62] K. Ercikan, R. Arim, D. Law, J. Domene, F. Gagnon, and S. Lacroix, "Application of think aloud protocols for examining and confirming sources of differential item functioning identified by expert reviews," *Educational Measurement: Issues and Practice*, vol. 29, no. 2, pp. 24–35, 2010.
- [63] G. Bonnet, F. Daems, C. de Clopper et al., *Culturally Balanced Assessment of Reading (c-Bar) A European Project*, European Network of Policy Makers for the Evaluation of Education Systems, Paris, France, 2003.

