

Translation, Linguistic Validation, and Electronic Administration in the Indian Languages: A Case Study Highlighting the Special Difficulties that Exist at this Intersection

Melissa Handa, MA; Barbara Brandt, MA; Mary C. Gawlicki, MBA; Shawn McKown, MA; Naoto Hasegawa, BA
— Corporate Translations, Inc. —

OBJECTIVE

This poster seeks to enumerate the technical challenges and subject comprehension issues associated with the intersection of Indian languages, linguistic validation, and electronic administration and recommends the use of a specialized list of precautionary measures to avoid escalating potential difficulties.

BACKGROUND

Translation and linguistic validation of questionnaires for either paper or electronic administration require similar procedures. Both methods of administration share the common difficulties of how to best utilize existing processes to accommodate Indian languages. Working with Indian language instruments can be challenging, even without the compounded effect of managing both ePRO and linguistic validation together.

METHODS

A case study was performed to compare and contrast the complexity of the processes used to linguistically validate an instrument for ePRO and paper administration. A fibromyalgia e-diary and corresponding paper diary were translated into four Indian languages and three non-Indian languages with varying levels of linguistic difficulty. This situation allowed for a multidimensional analysis of issues faced in translation for ePRO versus paper administration, and translation for Indian versus non-Indian settings.

ePRO, TRANSLATION, and LANGUAGE-SPECIFIC DIFFICULTIES

During linguistic validation, difficulties arise pertaining to font and formatting requirements, localization of colloquial terms and phrases, and maintaining concept equivalency while respecting cultural appropriateness. Additionally, preparing an ePRO instrument for use involves thorough analysis of existing instrument versions and possible modification of instructions, response sets, or questionnaire format to accommodate technology or screen size limitations. When preparing an Indian language for electronic administration, one must take even more care, taking into consideration the factors listed above as well as performing font-testing on the ePRO platform prior to production, and including additional proofreading steps to ensure accurate file conversion. The following is a list of common difficulties faced when preparing a translation for electronic administration—the extent to which each category must be considered depends on both the language and the format of administration:

- Formatting:** Text effects required for emphasis may not be supported by electronic format, such as italics, boldface, or underlining, which are necessary alternatives to capitalization in many languages. Consensus between translation resource and ePRO vendor should be reached at the outset of the project.
- Special Characters:** Special characters may not be supported by the necessary software. In extreme cases, a laminated card with term equivalency may be distributed to respondents to accompany the electronic questionnaire.
- Diacritics:** Diacritics (a mark added to a letter to indicate a special pronunciation) may initially be stripped during migration to electronic format, requiring fixes and multiple rounds of proofreading.
- Screen Space:** Localized source text may extend beyond allowable characters per screen limit, requiring additional screens.
- Rewording:** When producing a translation from English, it may not be possible to mirror question word order or structure exactly. As a result of screen size constraints, complete sentences rather than segments of sentences may need to be repeated in the target language when an instrument is used on an ePRO device.
- Font Compatibility:** ePRO software may not support some fonts, especially in the case of Indic and Dravidian language families. Some characters may not appear clearly. Solutions to this problem should be explored prior to beginning the project if this situation is likely.

A CASE STUDY: TRANSLATING AN E-DIARY AND ITS PAPER COUNTERPART

To explore the categories of comprehension difficulties and their prevalence in both Indian and non-Indian settings, a case study involving an eDiary and its paper counterpart was utilized. Linguistic validation was performed according to the special list of precautions for translating Indian languages and adapting for e-administration as outlined in **Figure 1**.

To preempt difficulties associated with Indian language translation, a concept elaboration guide was created by a survey research expert and utilized by translators and project management for complex terminology within the questionnaires; back-translations were reviewed by a survey research expert versed in the cultural requirements and linguistic structure of Indian languages; the same expert assisted with the communication between the US-English speaking project management and the Indian-English speaking translators; and interviewers clearly informed subjects that there were no right or wrong answers and that their interpretation of the material was what was being assessed.

To accommodate special issues related to ePRO, our project team ensured that the Excel spreadsheet used for the translation of the ePRO questionnaire could accommodate Urdu's right-to-left text orientation, tested fonts to identify the best ones for use with the ePRO vendor's character limit per screen, proofread screenshots to ensure proper display, implemented necessary changes resulting from cognitive debriefing in an Excel spreadsheet, color-coded the new text for ease of modification, and proofread modified screenshots before finalizing the translations.

Even after implementation of these precautionary measures, subject debriefing was crucial in order to identify and correct added difficulties that Indian subjects face when dealing with the technical terminology and complex instructions associated with e-administration. The eDiary and the Back-up Paper Diary were translated and debriefed within nine distinct cultural/linguistic settings, four of which were Indian languages (two from the Indic language family and two Dravidian). The demographic and summary data displayed in **Table 1** showcase comparable results between the electronic and paper questionnaires and between language groups.

Table 1. Demographic characteristics of the subjects by target language*

	Demographics			Item Comprehension	
	Residence (Cities)*	Age (Years)	Gender (M:F)	Education (Years)	eDiary
Afrikaans South Africa (5)	18-63	2:3	5-15	99.60%	99.40%
Chinese Taiwan (3)	24-64	2:3	11-14	100.00%	99.40%
Czech Czech Republic (4)	32-75	2:3	10-18	92.60%	99.40%
English UK (3)	18-49	2:3	11-18	100.00%	100.00%
Spanish-1* US (4)	26-58	2:3	9-18	100.00%	---
Spanish-2* US (5)	20-56	2:3	11-18	---	97.10%
Hindi India (4)	19-54	3:2	10-15	99.60%	100.00%
Kannada India (5)	22-55	3:2	10-12	100.00%	95.40%
Telugu India (4)	27-65	3:2	5-10	100.00%	95.40%
Urdu India (4)	26-62	3:2	7-17	98.70%	95.40%
Overall	19-65	12:13	5-18	98.94%	97.94%

* n=5 for each language
** For Spanish, countries of origin
*** The first subset of Spanish subjects tested the e-Diary, the second tested the Paper Diary

Tables 2 and 3 present comprehension issues raised by Indian subjects with regard to the e-Diary and Paper Diary. Almost all issues corresponded to instructions (which varied greatly according to mode of administration), rather than actual survey items or response scales. Subjects had difficulty understanding a greater number of items in the e-Diary; however, issues within the Paper Diary affected more subjects—causing a slightly lower overall comprehension rate. Most issues with both questionnaires were related to the use of technical or technological terminology (such as fibromyalgia, pin code, or training mode) or from complex instructions.

In three of four Indian languages, subjects faced confusion in relation to the instruction that the Paper Diary was to be used only in the event that the subject was not able to access the e-Diary via phone. To correct this, the title was transliterated for Telugu speakers; for Kannada, the title was replaced with the phrase "Additional daily diary used so that information is not lost: Instructions for use." While Hindi speakers did not have trouble with this concept, they found the use of the phrase "training" in the e-Diary to be ambiguous, confusing it with "education"—the term "practice" was used to replace this unclear term.

Table 2. Subject Comprehension Issues—eDiary

Problem	Subjects Affected	Languages Exhibiting Difficulty
The PIN codes do not match, please try again	1	Hindi
Hello, please complete your enrollment for the questionnaire by entering a 4 digit PIN code below	1	Urdu
Welcome, please click on the link below to complete today's questionnaire	1	Urdu
Select the number that best describes your fibromyalgia pain in the past 24 hours	1	Urdu
You do not need to complete this questionnaire today	1	Urdu
If you would like to complete the training questionnaire again, click on the link below	3	Hindi, Urdu
Thank you for completing the training questionnaire today, no data was stored	3	Hindi, Urdu
Your enrollment has been successful. To complete the questionnaire in training mode please use the link below	3	Hindi, Urdu
No data will be collected during training mode	3	Hindi, Kannada
Your next & future text messages will be a different link to the live questionnaire, please follow the same steps as training for the remainder of the study	2	Hindi
Restart Training	2	Hindi
Please follow the link below to complete your enrollment	1	Urdu

Table 3. Subject Comprehension Issues—Paper Diary

Problem	Subjects Affected	Languages Exhibiting Difficulty
Back Up Daily Diary: INSTRUCTIONS FOR USE	8	Kannada, Telugu
Please complete your daily diary using this paper booklet ONLY WHEN YOU ARE UNABLE TO USE THE PHONE.	7	Kannada, Telugu
Complete this diary every day in the afternoon/evening after your daily activities until you are able to respond by phone again	5	Urdu
Complete a separate page for each day you need to complete the paper diary	3	Urdu
dd mmm yyyy	1	Kannada
Version 1.0	5	Telugu

Although less than half of the subjects were from Indian language backgrounds (44%), 84.6% of all e-Diary comprehension issues were raised by Indian subjects and 78.4% of all Paper Diary issues. Far more comprehension difficulties were raised by speakers of the Indic languages, Hindi and Urdu, than by speakers of the Dravidian languages, Kannada and Telugu—especially with regard to the eDiary.

This discrepancy may result from the vast disparities between social classes within the Indian subcontinent, which cause subjects with lower educational attainment to have limited or no access to e-survey technology, computers, or cell phones in general. Since English is the default language within businesses and schools, subjects with greater educational attainment may also be unfamiliar with this type of technological terminology when it is rendered in an Indian language—in many instances, it was necessary to transliterate the English terms in order to adequately convey the concepts.

CONCLUSIONS

Linguistic validation of ePRO questionnaires for use in Indian settings may present special technical and comprehension challenges as compared to other language families. Indian localization for both paper PRO and ePRO requires additional considerations for special fonts and cultural appropriateness; it also faces satisficing and low item comprehension during linguistic validation via both modes. Evidence suggests that additional consideration must be exercised early on in order to avoid compounded difficulties and delays later in the process. The cognitive debriefing phase is a complementary and critical tool to aid in identifying comprehension problems related to the requisite technical terminology employed in the instructions associated with such devices. The use of a specialized list of precautions at the outset of the project is recommended to assist in preempting future difficulties.

REFERENCES

- Gawlicki M. Procedures for Linguistic Validation of Health Status Questionnaires. Mansfield Center, CT: Corporate Translations, Inc. 2004.
- Leveling language barriers (n.d.). Retrieved March 6, 2009 from, http://www.drivenbydefinition.com/html/linguistic_validation.htm
- Wild D et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. Value in Health Vol. 8 No. 2, 2005.
- Wild D et al. Multinational Trials—Recommendations on the Translations Required, Approaches to Using the Same Languages in Different Countries, and the Approaches to Support Pooling the Data: The ISPOR Patient-Reported Outcomes Translation and Linguistic Validation Good Research Practices Task Force Report. Value in Health Vol. 12 No. 4, 2009.
- Tiplady B. Electronic Patient Diaries and Questionnaires: ePRO Now Delivering on the Promise? The Patient: Patient-Centered Outcomes Research 2010;3(3):179-183.
- The CIA World Factbook, <https://www.cia.gov/library/publications/the-world-factbook/geos/in.html>, 23 Sep 2010.
- Sridhar KK. Language in Education: Minorities and Multilingualism in India. International Review of Education 1996; 42(4):327-47.
- Bhatt RM. World Englishes. Annual Review of Anthropology 2001;30:527-50.
- Census of India (n.d.). C-18: Population by Bilingualism, Trilingualism, Age and Sex. Retrieved 23 Sep 2010 from censusindia.gov.in/Tables_Published/C-Series/C.../DDWC-000018.pdf

Figure 1. Special Considerations for Indian Language Translation and ePRO Adaptation

