

# Gender Differences in Cognitive Debriefing of Translated Patient Questionnaires

Mary C. Gawlicki, MBA; Barbara A. Brandt, MA; Shawn McKown, MA; Carolyn Schulz, BA; Matthew Talbert, MA

— Corporate Translations, Inc. East Hartford, CT, USA —

## OBJECTIVE

The objective of this study was to determine if any discernible differences, in quantity or content, exist in feedback from male and female subjects provided during cognitive debriefing interviews of translated patient questionnaires. After a Clinical Outcomes Assessment (COA) questionnaire is translated, the final harmonized translation is pilot-tested on a sample of 5 subjects in the target language and country. For certain patient populations, such as Lupus patients, one gender may be affected more than the other, resulting in an overrepresentation in a patient sample group recruited for cognitive debriefing interviews.

## BACKGROUND

Existing literature shows that there are some differences between males and females in terms of participation and questionnaire response in some settings. Concerning participation, there is evidence that women may be less likely to participate in a focus group or classroom discussion when in the presence of men. Conversely, when in the presence of all females, there is a likelihood of increased participation and contribution [1]. In the context of the online survey environment, a recent study investigating online survey participation showed that females are underrepresented. While the authors controlled for several variables that may influence participation, such as age and technological ability, it was determined that males are more inclined to participate in an online environment [2]. A study regarding sensitive topics found that the subject matter of a questionnaire may elicit different responses from males or females. A study by Sikweyiya and Jewkes showed that female participation lacked in a survey on domestic violence, as a result of fear of further abuse or violence from their male domestic partner for having participated [3]. Another such example, specific to linguistic validation, concerns cognitive debriefing of sensitive topics. A study by Gawlicki et. al. showed that men may be less inclined to participate in cognitive debriefing interviews on erectile dysfunction if the interviewer was female [4]. However, no research has been conducted to-date to determine the content and quality of males' and females' responses in a linguistic validation setting. The current study seeks to determine what gender differences may exist in terms of responses and contributions to a final translated COA.

The following examples demonstrate how comments from cognitive debriefing interviews are categorized:

**Suggested revisions to improve conceptual clarity:** Subjects will identify concepts that are difficult to understand and will suggest revisions to improve the conceptual clarity of the source, resulting in a translation change.

Example: Suggested revision to improve conceptual clarity	
Original text	Revised text
Please answer the following items	Please answer the following questions

Although “items” may be appropriate in many languages such as Spanish, French and Italian, the concept was not appropriate in some languages, such as Korean and Greek.

**Suggested revisions to improve cultural appropriateness:** Subjects may identify certain items that are not appropriate for their language or country. If several subjects take issue with a particular item, even if they understand the original concept, a change to the translation is often made. The following revision example was made as a result of debriefing for English-UK.

Example: Suggested revision to improve cultural appropriateness	
Original text	Revised text
It is difficult to put on pants	It is difficult to put on trousers

This revision was made because “trousers” is a necessary adaptation for the United Kingdom, although the term “pants” is conceptually equivalent.

**Suggested revisions that are stylistic:** Some comments made by subjects may not lead to a translation revision, on the grounds that the revision suggestion is stylistic. If a subject or subjects make a suggestion to revise the translation as a conceptually equivalent term based on preference alone, despite understanding the original translation, the revision will most likely not be made.

Example: Stylistic suggested revision	
Original text	Rejected revision suggestion
There is pain in my belly	There is pain in my abdomen

This revision was rejected as “abdomen” is conceptually equivalent to the source concept, “belly,” and does not have any impact on improvement of comprehension.

**Suggested revisions that deviate from the source text:** Subjects may also make suggestions which are deviations from the source text.

Example: Source text deviation suggested revision	
Original text	Rejected revision suggestion
I have trouble doing household activities	I have trouble cleaning the house

This revision was rejected on the grounds that “household activities” encompasses activities other than “cleaning the house” and therefore, the suggested revision is a source text deviation.

Subjects also made comments unrelated to the translation and that could not result in a change to the translation. Such comments were classified as “miscellaneous.” Examples of such comments are: “I cannot do household chores anymore because of my disease” in reference to a physical assessment of household tasks, or “My house doesn't have stairs” when asked about their ability to use stairs.

## METHODS

The cognitive debriefing results of 4 disease-specific questionnaires were analyzed:

- Chronic Obstructive Pulmonary Disorder (COPD) questionnaire– 790 words
- Stroke questionnaire– 769 words
- Diabetes questionnaire– 523 words
- Ulcerative Colitis questionnaire– 888 words

54 different languages were represented in the sample.

The following was tabulated during analysis of data collection forms (DCF):

- Comments made overall per males and females
- All comments were then categorized by the following:
  - Suggestions that led to a translation revision
    - Conceptual Clarity
    - Cultural Appropriateness
    - Grammatical
  - Suggestions that were stylistic
  - Suggestions that deviated from the concept in the source text
  - Miscellaneous comments
- The mean was calculated for males and females in all categories
- A two-tailed t-test was completed to determine whether the difference in means between male and female comments in each category was statistically significant
- Age and years of academic education of each subject were also tabulated

## RESULTS

Table 1: Aggregated means of all debriefed questionnaires per DCF (n=351)								
	n	Age	Education	TOTAL comments	Suggestions that improved the translation	Stylistic suggestions	Source deviation suggestions	No revision suggestion
Total males	171	55.1	12.7	3.1	0.75	0.91	0.71	0.67
Total females	180	51.7	12.6	4.2	0.91	1.55	0.89	0.88

Table 2: Aggregate p-values	
p-value for total comments made	0.006
p-value for suggestions that improved the translation	0.162
p-value for stylistic suggestions	0.021
p-value for suggestions that deviate from the source text	0.227

Table 3: Breakdown of types of revisions made		
	Total	% of total changes
Females-Conceptual Clarity	128	77.5%
Females-Cultural Appropriateness	15	9%
Females-Grammatical	22	13.5%
Females- Total suggestions leading to a revision	165	100%
Males-Conceptual Clarity	100	78.1%
Males-Cultural Appropriateness	12	9.4%
Males-Grammatical	16	12.5%
Males- Total suggestions leading to a revision	128	100%

## CONCLUSIONS

The results show the following: females averaged 4.2 comments per DCF, and males averaged 3.1 comments per DCF. For total comments made, the p-value was 0.006, which shows a statistically significant difference between one's gender and the number of overall comments made per DCF. The p-value also shows more evidence to support the hypothesis that females make more comments per DCF overall.

Because the end goal of linguistic validation is to ensure that the translated text is comprehensible and culturally appropriate, the primary focus of the present research was to analyze suggestions that led to a revision of the translation versus those that did not result in a revision. Knowing whether or not there is a difference between male and female responses in this specific category is important, while stylistic suggestions and suggested source deviations do not contribute to the end product. Our research showed that females make an average of 0.91 comments per DCF that result in an improvement to the translation, while males make 0.75 comments that result in an improvement to the translation. Although it is clear that females make more comments, the difference of 0.16 is marginal. Moreover, the p-value for comments that result in an improvement to the translation is 0.162, showing no meaningful difference between male and female contributions to the final translation product. Additionally, as demonstrated in Table 3, there is no difference between males and females in terms of the kind of revision that is made as a result of their suggestion.

To conclude, there are some differences between male and female contributions to cognitive debriefing interviews in that females make more comments overall, specifically stylistic comments. However, in terms of meaningful contribution to the final translation, there is no significant difference between men's and women's contributions in making translations more comprehensible, culturally appropriate and grammatically correct.

## REFERENCES

1. Brau, Julia, Paayal Desai, Akmaral Omarova, Jung Paik, and Julie Sandler. “Gender Discrepancies in Academic Performance.” The Harbus. Harvard Business School, 5 May 2010. Web. 18 Apr. 2014.
2. Smith, William G., Ph.D. “Does Gender Influence Online Survey Participation?: A Record-linkage Analysis of University Faculty Online Survey Response Behavior.” ERIC - Education Resources Information Center, 6 June 2008. Web. 18 Apr. 2014.
3. Sikweyiya, Yandisa, Rachel Jewkes, and Jeremy Miles. “Perceptions and Experiences of Research Participants on Gender-Based Violence Community Based Survey: Implications for Ethical Guidelines.” PLoS ONE 7.4 (2012): E35495.
4. Gawlicki, M., Handa, M., and McKown, S. “Special Precautions to Consider When Performing Cognitive Debriefing of Sensitive Topics.” Value in Health, May 2011. Vol. 14, Issue 3, Page A113.