

Genre Misrecognition in Practice: Bridging the Theory-Method Gap in Iranian Questionnaire Translation¹

Nasimeh Nouhi Jadesi², Marziyeh Sadeghzadeh³
Narjes Nasrpour^f & Fariba Bozarjomehri⁵

Abstract

Translation of psychometric instruments represents a specialized translation genre where the questionnaire serves as a calibrated stimulus-response tool. Within the framework of Translation Studies, the success of such transfers is measured by functional equivalence and the preservation of metric invariance rather than mere formal correspondence. This study evaluates the methodological approaches to the translation phase in Iranian psychology and education research by analyzing the type and frequency of translation guidelines cited in instrument validation studies. Using a descriptive-analytical design, we examined a corpus of 451 studies authored by Iranian researchers appearing in Iranian psychology and education journals. The findings revealed that only 3.8% of domestic studies explicitly identified a translation protocol. Furthermore, domestic articles remain heavily anchored in Brislin (1970) as their primary guideline, a dated paradigm, favoring back-translation as the principal method of questionnaire transfer and prioritizing mechanical reversibility over holistic, team-based adaptation. This methodological stagnation suggests a deeper conceptual issue: these patterns can be interpreted as evidence of genre misrecognition, as questionnaires are treated primarily as linguistic texts rather than measurement instruments requiring specialized translation and adaptation procedures. We argue that overcoming this gap requires the institutionalization of genre-specific standards: multi-step translation procedures, cultural adaptation, and expert-led adjudication.

Keywords: Cross-cultural survey translation, Iranian researchers, Psychology and education articles, Questionnaire translation, Questionnaire translation guideline

1. This paper was received on 15.12.2025 and approved on 17.01.2026.

2. Assistant Professor, Department of English, Faculty of Humanities, Salman Farsi University of Kazerun, Kazerun, Iran; email: dr.noohi@kazerunsfu.ac.ir

3. Corresponding Author: Associate Professor. Department of Psychology, Faculty of Humanities, Salman Farsi University of Kazerun. Kazerun, Iran; email: sadeghzadeh@kazerunsfu.ac.ir

4. M.A., Department of English, Faculty of Humanities, Salman Farsi University of Kazerun, Kazerun, Iran; email: narjesnasrpour@gmail.com

5. M.A., Department of Psychology, Faculty of Humanities, Salman Farsi University of Kazerun, Kazerun, Iran; email: faribabozarjomehri@gmail.com

Introduction

In the era of globalized psychological research, the translation and cross-cultural adaptation of questionnaires as measurement instruments has become a fundamental methodological requirement. In Iran, reliance on psychometric scales originally developed in Western contexts has increased significantly, making the translation phase a critical bridge between cultural and linguistic realities (Sadeghzadeh & Nouhi Jadesi, 2026). In Translation Studies, a questionnaire does not fit neatly into Newmark's (1988) categories of **expressive** texts (author-centered, stylistic focus) or **informative** texts (fact-centered, clarity-oriented); rather, it functions as a calibrated instrument for measurement. As elaborated by Repke and Dorer (2021), questionnaires represent a unique text type and serve as measurement tools that must operate effectively. They should accurately gauge the intended information and ideally reflect the original questionnaire's meaning consistently across all language versions. Consequently, any error or oversight during the translation process does not merely result in stylistic awkwardness but also leads to measurement bias, which is defined as distortions that occur when translated items do not capture the same underlying construct as the original instrument, leading respondents in different language versions to interpret or answer questions differently or undermining the comparability and validity of survey data across languages and cultures (Behr et al., 2016; Van de Vijver & Tanzer, 2004). It must be admitted that various factors- such as limited resources, journal word-count restrictions, and prevailing disciplinary practices- may also lead to incomplete reporting of rigorous protocols. However, this bias is particularly dangerous because, as pointed out by Heggstad et al. (2019), if the initial translation of an instrument is flawed, every subsequent piece of data collected and every clinical conclusion drawn from the translated questionnaire becomes scientifically questionable, a phenomenon called the cascading effect. Thus, a single translational error at the start can jeopardize years of empirical research.

The translation of questionnaires represents a distinct genre that differs fundamentally from other types of translation, such as literary or technical translation. From a Translation Studies perspective, the translation of psychometric instruments is a unique genre that demands a specific type of equivalence. Van de Vijver and Tanzer (2004) defined this as a multi-layered construct. It begins with semantic equivalence, ensuring words share meanings, but must progress to conceptual equivalence, which questions whether the psychological construct itself (e.g., anxiety or self-efficacy) exists in the same form within the target sociocultural landscape. While literary translation allows creativity to preserve the author's spirit, and technical translation focuses on the denotation of terms, questionnaire translation is strictly governed by the stimulus-response dyad. As argued by Harkness et al. (2003), the translator must achieve

functional equivalence, meaning that they must ensure that the target version of an item triggers the same psychological response as the source version. In this genre, the translator must respect metric invariance. For instance, a slight change in the intensity of a frequency adverb (e.g., sometimes vs. occasionally) can alter the statistical properties of the scale. This requires a level of precision and specialized knowledge that goes beyond general translation competence. This positions the translator not merely as a linguistic conduit, but as a cultural mediator and methodological co-researcher, a role that demands a structured, accountable process (Repke & Dorer, 2021). Recognizing these complexities, several rigorous quality control mechanisms have been put in place.

One such mechanism is using questionnaire translation guidelines. Various research methodologies and translation experts (e.g., Beaton et al., 2000; Brislin, 1970; Wild et al., 2005), along with some organizations (e.g., World Health Organization (WHO, 2016); International Testing Commission (ITC, 2017), have developed standardized questionnaire translation guidelines. They provide translators and researchers with a step-by-step roadmap to ensure that the final target version is culturally and psychologically equivalent to the source version, and to minimize subjectivity and provide a transparent framework for the adaptation process.

However, the extent to which these internationally recognized guidelines are reflected in the reporting and methodological practices of questionnaire translation in Iranian psychology and education research remains largely unexplored. Despite the critical importance of these quality control measures, the actual practice among Iranian researchers in terms of choosing and using questionnaire translation guidelines remains under-researched. While many Iranian studies claim to follow standard procedures, there is a lack of transparency regarding which specific guidelines are actually employed and the justification for their choices. *This study, therefore, addresses the question: What are the specific practices in psychology and education, two fields with a particularly high reliance on translated psychometric instruments? The present study, therefore, aims to fill this gap by investigating which questionnaire translation guidelines are most frequently used by Iranian researchers in the fields of psychology and education. Accordingly, this study asks: What translation guidelines are reported in questionnaire validation studies conducted by Iranian researchers in psychology and education, and how frequently are these guidelines used?*

Literature Review

The early stages of questionnaire translation research were dominated by a quest for linguistic correspondence, a paradigm largely defined by the work of Brislin (1970). This

model introduced back-translation as the primary yardstick for accuracy, under the assumption that if a text could be translated into a target language and then translated back to the source without losing its original wording, the translation was satisfactory. However, as Acquadro et al. (2008) highlight, this literalist approach was increasingly challenged for its failure to capture cultural nuances. Critics argued that back-translation often creates versions that are linguistically correct but psychologically hollow in the target culture (Behr & Braun, 2023) or as put by McKenna and Doward (2005), “back translation has no clear scientific basis and its use casts doubts on the ability of translators” (p.89).

A pivotal shift occurred in the 1990s as researchers began to prioritize the instrumental nature of these texts. Guillemin et al. (1993) provided one of the first comprehensive taxonomies of adaptation, explicitly distinguishing *cross-cultural adaptation* from mere linguistic translation. In their framework, adaptation denotes a broader process that integrates cultural, conceptual, and functional equivalence rather than language transfer alone. This evolution reached a significant milestone with the guidelines proposed by Beaton et al. (2000), which introduced a rigorous six-stage trajectory. This framework emphasizes that the psychometric properties of a tool- its reliability and validity- are inseparable from the cultural salience of its translated items. This shift in guidelines’ focus can be conceptualized along a spectrum from linguistic-centric to culture-centric models, a spectrum that mirrors the broader genre distinction between treating questionnaires as linguistic texts versus treating them as measurement instruments.

At one end, models like Brislin’s (1970) prioritize linguistic fidelity through back-translation. At the other end, comprehensive frameworks such as Beaton et al.’s (2000) embed translation within a multidisciplinary, iterative process of cross-cultural adaptation, emphasizing expert review and cognitive pretesting. The ISPOR Task Force principles for translation and cultural adaptation proposed by Wild et al. (2005), together with the International Test Commission (ITC) Guidelines (2017), further institutionalize this rigorous, team-based approach. Harkness (2003) proposed TRAPD (Translation, Review, Adjudication, Pretest, and Documentation), which is now considered the gold standard for questionnaire translation.

Despite the well-established global frameworks for instrument adaptation, a significant empirical gap remains regarding the actual implementation of these quality control mechanisms within many academic landscapes. A systematic review by Bowden and Fox-Rushby (2003) evaluating translation practices across Africa, Asia, Eastern Europe, the Middle East, and South America similarly found significant shortcomings in establishing functional equivalence, concluding that broader research practices and adherence to translation

guidelines in these regions still require substantial improvement. While the international literature, ranging from the foundational stages of Beaton et al. (2000) to the modern team-based imperatives of TRAPD (Harkness et al., 2003, 2010), emphasizes a rigorous multi-step trajectory, it is unclear to what extent researchers adhere to these standards. Similarly, Maneesriwongul and Dixon (2004) reviewed 47 nursing studies to assess how instrument translation was conducted and reported. They found wide variation in practice, noting that most studies used forward-back translation but often omitted key steps such as testing, comparison of versions, or involving multiple translators. Their six-category classification demonstrates that translation procedures are unevenly applied and inconsistently documented, highlighting the need for clearer reporting standards and for empirical studies that map how translation guidelines and core procedures are actually reported in published research. Valdez et al. (2021) provide a conceptual taxonomy of translation frameworks in multilingual survey research, arguing for explicit disclosure of the approach used, emphasizing that guideline selection affects data quality and cultural validity. However, their contribution is normative rather than empirical: they do not assess whether specific guidelines (e.g., WHO, ITC) are actually named in published studies, nor do they quantify the reporting of core protocol components (e.g., dual translation, adjudication, cognitive interviewing). In contrast, the present study operationalizes this concern by systematically coding a corpus of 451 Iranian psychology and education articles for the explicit identification of named guidelines and the reporting of guideline-linked procedures, producing a frequency-based map of how translation standards are—or are not—documented in practice.

The current domestic discourse lacks a systematic and critical evaluation of whether foundational dated guidelines in favor of back translation or more recent team-based guidelines are being employed in Persian validation studies published in Iranian journals. This study addresses this niche by providing the first frequency-based mapping and critical review of the translation guidelines utilized in the articles published in Iranian psychology and education journals. The novelty of this research lies in its systematic deconstruction of the Iranian status quo. By mapping the frequency of specific translation guidelines, this study moves beyond theoretical discussion to reveal the practical habits and potential methodological shortcuts prevalent in the field. Identifying these common practices is essential for understanding the specific risks of measurement bias inherent in the translated version of research instruments.

Method

The present study utilizes a systematic review and descriptive-analytical design to investigate the translation guidelines adopted by Iranian researchers in their questionnaire

translation studies. The research was conducted in three distinct rigorous processes: corpus construction, **procedural** mapping, and critical analysis.

This study developed and analyzed a dataset comprising studies in psychology and education that focused on questionnaire translation and validation published in Iranian domestic journals. To this end, we included studies whose final reports were published in either Persian or English, provided that the study involved translating an instrument into Persian, in accordance with the inclusion criteria.

The sampling frame for this study was purposively restricted to ensure methodological coherence and analytical rigor. Specifically, we limited the dataset to studies published in Iranian domestic journals that reported translation and validation procedures, as this focus allowed for consistent comparison of practices within a defined scholarly context. The researchers utilized the official list of peer-reviewed journals approved by the Ministry of Science, Research and Technology (MSRT). Within this elite tier of publications, specific journals dedicated to the psychology and education fields were targeted (N = 451). This was to ensure that the **article** corpus compiled is rigorous enough and had undergone a rigorous peer-review process before publication. This study focuses exclusively on Iranian-published psychology and education journals. As a result, it does not capture questionnaire translation practices reported by Iranian authors publishing in international outlets, where methodological expectations may differ. Future research could extend this mapping to include internationally published work, providing a more comprehensive picture of Iranian researchers' translation practices across publication contexts.

A methodological limitation of this study is its focus on articles published in elite-tier Iranian journals. Because these outlets typically apply stricter editorial and methodological standards, the patterns observed here may represent a relatively higher level of reporting quality. Translation practices in lower-tier or less-regulated journals may be less rigorous and potentially reveal an even wider gap in protocol adherence. Consequently, the present findings should be interpreted as a conservative estimate of the problem rather than a comprehensive reflection of all questionnaire translation practices in Iran. The search was conducted manually by accessing the individual archives of each journal and utilizing their internal search engines. The manual search was conducted between September 15 and December 10, 2025. To ensure a comprehensive dataset, both Persian and English-language articles were included. The search queries employed a combination of Boolean operators and specific keywords: "questionnaire translation," "scale," "validation," "psychometric," and "Persian version," along with their Persian equivalents (e.g., روان‌سنجی, هنجاریابی, and ترجمه پرسشنامه).

A time frame of 2000–2025 was set, which allowed for a comprehensive view of the evolving translational practices within these Iranian publications. The resulting corpus was refined through a strict screening process. The inclusion criteria were limited to studies that involved the translation and cross-cultural adaptation of existing English-language instruments into Persian, followed by psychometric validation. To strengthen transparency, we also applied explicit exclusion criteria: studies reporting the *de novo* development of original questionnaires; studies examining questionnaire responsiveness or clinical/health outcomes without a primary focus on the translation phase.

Having compiled the dataset, each article was scrutinized to determine if a formal translation guideline was explicitly reported. Two researchers screened and coded the corpus. Coding was not conducted independently because each researcher was assigned half of the articles. The coding task involved only binary identification of whether an article explicitly cited a translation guideline (present/absent), a highly objective judgment for which formal inter-rater statistics such as Cohen's kappa are not typically informative. To ensure consistency, the researchers jointly cross-checked a subset of articles and discussed any uncertain cases until agreement was reached. A study was considered to have 'reported' a guideline if it explicitly named a recognized framework (e.g., Beaton, TRAPD, ITC) or cited the source (Beaton et al., 2000). The decision to count only explicit citations reflects a methodological boundary of the study. Procedural descriptions, such as back-translation, are too generic to infer which guideline the authors relied on, given that this step appears across several frameworks and is often transmitted informally in local academic practice. Coding them as instances of Brislin or any specific guideline would introduce inference error. Restricting the analysis to reported citations, therefore, preserves coding validity and transparency. A frequency analysis was then performed to provide a quantitative overview of which guidelines Iranian scholars most frequently rely upon as their primary quality control mechanism.

Because this study analyzed publicly available journal articles and did not involve human subjects, no ethics approval was required. All journal archives were accessed through publicly accessible platforms or institutional subscriptions in accordance with standard scholarly practice.

Results

The systematic analysis of questionnaire translation studies published in Iranian journals revealed a rarity in terms of reporting questionnaire translation guidelines used in the articles published by Iranian researchers in psychology and education, and published in Iranian journals. Within the corpus of 451 studies published in Iranian national journals, only

twenty studies (3.8%, on average) explicitly cited a formal translation guideline. It indicates that the translation process itself remains largely undertheorized in published research related to these disciplines, where the methodological foundation of instrument translation is overwhelmingly left unreported. The tables below present the obtained results.

Table 1. *Summary of Screening and Reporting of questionnaire translation studies published in Iranian journals in the psychology and education fields*

| Discipline | Total Articles Screened | Articles Reporting Guidelines | Reporting Rate (%) |
|------------|-------------------------|-------------------------------|--------------------|
| Psychology | 360 | 17 | 4.4% |
| Education | 91 | 3 | 3.2% |
| | | | 3.8% on average |

More critically, the analysis of the types of guidelines employed reveals a profound methodological stagnation, reflecting a continued neglect of genre-specific adaptation standards required for psychometric instruments. This stagnation indicates that translation practices remain confined to linguistic reversibility frameworks rather than functional, genre-appropriate equivalence. As shown in Table 2, Iranian researchers employed a narrow range of guidelines; specifically, only six different frameworks were utilized.

Table 2. *The Type and Percentage of Questionnaire Translation Guidelines Reported by Iranian Researchers in Iranian Journals*

| Translation guideline use | Frequency | Percentage |
|-------------------------------------|-----------|------------|
| Brislin (1970) | 9 | 45% |
| Beaton et al. (2000) | 3 | 15% |
| Jones et al. (2001) | 3 | 15% |
| Sousa & Rojjanasrirat (2011) | 3 | 15% |
| Int. Quality of Life (IQOLA) (2008) | 1 | 5% |
| Chen & Boore (2010) | 1 | 5% |
| Total Reporting | 20 | 100 |

The landscape within domestic publications was characterized by extreme restriction. The small subset of reporting studies ($n = 20$) relied on only six different guidelines. Brislin (1970) dominated this space, representing 45% ($n = 9$) of reported protocols. The continued dominance of Brislin's model, despite the availability of more contemporary protocols, likely

reflects several factors: institutional path dependency, the perception that it is simple and efficient to implement, and the widespread misconception that a successful back-translation is equivalent to demonstrating validity. The remaining citations were distributed between Beaton et al. (2000), Jones et al. (2001), and Sousa and Rojjanasrirat (2011) with an equal rate of 15%, and Acquadro et al. (2008), the International Quality of Life (IQOLA) protocol, and Chen and Boore (2010), each having been used only once.

Notably, comprehensive, contemporary frameworks that define current best practice in the translation of measurement instruments—such as the team-based TRAPD model (Harkness, 2003) or the ITC Guidelines (2017)—were absent from the corpus. This absence highlights a critical disconnect between local translation practices and the evolving theoretical and methodological standards of the Translation Studies field as applied to instrument translation.

Discussion

The findings of this review reveal a fundamental genre misrecognition—understood here as the tendency to treat psychometric instruments as ordinary linguistic texts rather than as a distinct functional genre requiring specific adaptation procedures—within Iranian translation practice. The near-invisibility of formal translation protocols and the hegemony of the Brislin (1970) model demonstrated that questionnaires are routinely treated as standard informative texts rather than as calibrated stimulus-response instruments that define their unique genre and its treatment in the Iranian academic practice. The persistent dominance of the Brislin (1970) model serves as the primary empirical indicator of this misrecognition.

As Behr (2018) clearly elaborates, a questionnaire is a specific genre, and its translation requires competence beyond the general translation competence expected from translators. Taken together, they directly support the study's stated contribution: this is the first frequency-based empirical mapping of guideline use within Iranian questionnaire translation research. The notably low rates of explicit reporting—such as the predominant reliance on Brislin, the absence of WHO guidance, and the near-absence of culture-centric procedures—provide the empirical foundation for the study's critical deconstruction of current practice. These patterns make visible the methodological assumptions representing the Iranian status quo, thereby substantiating the novelty claim introduced in the Introduction section above.

The overwhelming dominance of Brislin's (1970) back-translation model, accounting for 45% of the scant guidelines reported, is particularly alarming from a Translation Studies standpoint. It is theoretically inappropriate for questionnaires as a unique genre, as Brislin's model verifies only lexical reversibility, whereas the questionnaire genre requires functional

equivalence—ensuring that a translated item elicits the same psychological meaning and respondent-level effect as the original (Harkness et al., 2003).

This model operates on a paradigm of semantic correspondence and lexical reversibility. It is designed to answer the question, "Can the words be translated back?" rather than the core question of instrument translation: "Does the translated item elicit the same cognitive and psychological response?" (Harkness, 2003). By optimizing for mechanical round-trip accuracy, the process risks producing translations that are linguistically "correct" but psychologically "hollow" or culturally misaligned (Behr & Braun, 2023), thereby failing to achieve the functional equivalence required for valid measurement. This is a systemic genre misrecognition, where the calibrated, stimulus-response instrument is handled as if it were a standard informative text, its translation rendered an insignificant, unregulated act.

The prevalence of back-translation in the domestic corpus is particularly concerning when viewed through the lens of recent experimental evidence. Behr and Braun (2023) demonstrated that back-translation is a 'limited' method for quality assurance, failing to rectify 22 out of 29 identified translation issues in their English-German case study. By contrast, team-based approaches resolved every issue. This suggests that the Iranian domestic research community is relying on a framework that is statistically likely to leave nearly 76% of translation errors, including mistranslations and wording issues, undetected in the final target translated version. They indicated that "while back translation can certainly pick up (gross) errors, it is less suitable to identify more nuanced differences, needed cultural adaptations or language-related issues such as ambiguity, awkwardness, grammatical, and syntactical problems, typical one-to-one equivalencies that may not fit for a given context, or inappropriate register, that is, the appropriate style for a given situation" (p. 288). Thus, the persistent use of Brislin's model is not merely a methodological choice but a symptom of the fundamental genre misrecognition identified in this review.

This methodological gap becomes starkly visible when comparing the structural components of different translation frameworks, as illustrated in Table 3. The table provides a comparative analysis of three representative guidelines—Brislin (1970), Beaton et al. (2000), and TRAPD (Harkness, 2003)—mapping their core philosophy and the procedural stages. This side-by-side comparison visually highlights the evolution from a linguistically-focused, individualistic model (Brislin) to comprehensive, team-based, and empirically-grounded protocols (Beaton et al., TRAPD). The absence of mandated stages for expert adjudication **and** cognitive pretesting in the Brislin model, which dominates Iranian practice, highlights the precise procedural shortcomings contributing to the genre misrecognition and measurement risks discussed below.

Table 3. A comparative representation of Cross-Cultural Adaptation Frameworks

| | | Stage 1 | Stage 2 | Stage 3 | Stage 4 | Stage 5 | Stage 6 | Stage 7 |
|--|---|--|---|--|---|--|---|--|
| Questionnaire translation guideline | Core Philosophy | Preparation | Initial translation | Synthesis / review | Back-translation | Adjudication/ committee | Pretesting | Documentation |
| Brislin (1970) | To achieve functional equivalence through an iterative, decentering process. Quality is evaluated by whether the source and target versions elicit similar responses. | Simplification of source text (no passive voice, no metaphors) | Forward Translation: One bilingual translator. | N/A | Blind Back-Translation: The second bilingual translates back to the source. | N/A | Multiple Criteria for Equivalence: E.g., 1) Bilingual administration (comparing responses), 2) Performance tasks, 3) "Random probe" interviews on item meaning. | N/A |
| Beaton et al. (2000) | A systematic, committee-driven guideline to maximize semantic, idiomatic, experiential, and conceptual equivalence. | Identify the adaptation scenario (e.g., same language/ different culture). | Dual Translation: Two independent translators (T1 & T2). | Synthesis: Translators and the observer produce the T-1,2 version. | Back-Translation: Two naive bilinguals translate T-1,2 to the source. | Expert Committee: Multidisciplinary team reviews all versions. | Pretesting: Cognitive testing on 30–40 subjects. | Submission: All reports are sent to developers for audit. |
| TRAPD (Harkness, 2003) | A team-based, modular approach emphasizing collaborative review and clear division of labor (T, R, A, P, D) to ensure quality and conceptual equivalence. | Team assembly and preparation of source materials. | Parallel Translation (T): Multiple independent translators. | Review (R): Team meeting to reconcile versions. | N/A | Adjudication (A): Senior expert makes final decisions. | Pretesting (P): Small-scale field test. Cognitive interviews with participants | Documentation (D): Explicit demand to document all stages. |

This isolation stands in direct contradiction to the contemporary view, established in the translation literature, of the translator as a cultural mediator and methodological co-researcher (Repke & Dorer, 2021). The current practice reduces the complex questionnaire translation process to a solitary linguistic exercise, devoid of the necessary interdisciplinary scrutiny. Contemporary gold standard models—such as TRAPD and the Beaton et al. (2000) framework—are fundamentally architectures for collaborative knowledge production. Similarly, the WHO (2016) guidelines and the ITC (2017) standards explicitly mandate multidisciplinary teamwork in which the initial translation is systematically reviewed, debated, and refined by additional translators, subject matter experts, and methodologists. The lack of mandated Review (R) and Adjudication (A) stages in the dominant local protocols eliminates this critical layer of dialogue. Consequently, decisions on crucial aspects—such as the cultural resonance of an idiom, the appropriate register for a target demographic, or the conceptual alignment of a psychological construct—are made in isolation, based on a single perspective, rather than being forged through the necessary collective negotiation. This isolation transforms a complex act of cultural mediation into a clerical, non-deliberative step, severing the translation process from the methodological rigor it requires. The modern translation theory, as exemplified by the TRAPD and Beaton et al. (2000) models, has moved toward a "calibrated stimulus" approach, ensuring that the translated item triggers the same psychological response and cognitive work in the target respondent as it did in the source population. By bypassing the Review and Adjudication stages inherent in modern models, Iranian researchers risk producing instruments that are linguistically accurate but psychologically inert.

In practice, this role manifests in several ways. First, translators make theoretically consequential decisions about item meaning, relevance, and cultural applicability; these decisions inevitably shape what the questionnaire measures in the target context. Second, they diagnose methodological problems embedded in the source instrument, such as idiomatic phrasing, culturally marked references, or conceptual gaps, and propose modifications that are necessary for maintaining construct validity. Third, translators participate in iterative team processes (e.g., adjudication, cognitive interviewing feedback, or committee harmonization), where their judgments directly influence item retention, rewording, or deletion. Finally, because they document rationales for choices and identify points of semantic risk, their input becomes part of the methodological transparency of the study.

The Neglect of Cognitive Verification and Empirical Grounding

A critical divergence highlighted by the comparative table is the role of pretesting and cognitive verification. In the older Brislin (1970) framework, verification is often reduced to a linguistic logic puzzle where the back-translation serves as the ultimate proof of quality. Conversely, Beaton et al. (2000) and TRAPD mandate cognitive pretesting as a fundamental stage to verify item interpretation within the target population. The absence of these stages in the domestic Iranian corpus suggests that researchers are substituting empirical data from real-world respondents with the subjective judgments of bilingual translators.

Epistemologically, the practical consequence of this methodological approach is a likely systematic compromise of conceptual (psychological) equivalence—the highest and most critical layer in Van de Vijver and Tanzer's (2004) multi-level equivalence framework, as it concerns whether the same underlying construct is being measured across cultures. If the psychological construct (e.g., "self-efficacy") is not faithfully evoked by the Persian items, the entire research enterprise built upon the translated instrument is undermined by a cascading effect of measurement bias (Heggstad et al., 2019). This is not a minor translational shortcoming but an epistemological flaw that jeopardizes the validity of locally produced knowledge. When construct equivalence remains unverified, Persian-language data cannot be confidently interpreted as measuring the intended psychological attribute. As a result, empirical findings derived from such instruments may not accurately reflect the constructs they claim to assess. Over time, clinical guidelines, research conclusions, and policy decisions built upon these data risk being grounded in measurements whose relationship to the intended constructs remains uncertain. In the Iranian context, where most instruments are imported through translation rather than developed locally, this problem reinforces the assumption of seamless knowledge transfer from the West to the East while overlooking the sociocultural specificities of the target community.

This study has several limitations. First, the search was restricted to MSRT-listed domestic journals, which may limit the generalizability of the findings to broader publication contexts. Second, although the search was conducted manually and across multiple sources, uncatalogued items may still have been missed. Third, the analysis relied on explicit citation of translation guidelines; therefore, unreported or implicit use of protocols could not be captured. Finally, the study did not assess the actual quality of the translations themselves, as its focus was limited to reporting practices rather than linguistic or psychometric evaluation.

Conclusion

The findings of this systematic review suggest a substantial gap between reported questionnaire translation practices in Iranian research and contemporary cross-cultural adaptation standards—a pattern that warrants further investigation into the underlying causes. This systematic review finds that only 3.8% of studies in the domestic corpus report using a formal translation guideline, and among those that do, Brislin's (1970) back-translation model is the most frequently cited (45%). These findings indicate a substantial gap between recommended practice and reported procedure. The observed pattern—low reporting rates and dominant reliance on a framework designed for lexical verification—raises questions about whether current practices align with the demands of the questionnaire genre, which requires cultural-conceptual adaptation. While scholars publishing internationally increasingly adopt multi-stage frameworks such as TRAPD, the domestic pattern may suggest that the specialized genre of questionnaire translation is often treated as a perfunctory linguistic exercise. Further investigation is required to identify the causes of low reporting rates and the persistent citation of Brislin's model. What can be concluded is that current reporting practices do not provide sufficient transparency to verify the functional equivalence of Persian-adapted instruments.

Moving forward requires a paradigm shift in translation practice. Researchers and translation experts in Iran must advocate for the institutional adoption of genre-appropriate frameworks such as TRAPD or the Beaton et al. (2000) guidelines. This shift entails:

- 1- Abandoning the back-translation loop as a standalone quality assurance tool.
- 2- Implementing mandatory, documented stages of expert adjudication and cognitive pre-testing.
- 3- Requiring journals and ethics committees to mandate transparent reporting of the translation protocol used.

Only by reclaiming questionnaire translation as a documented, collaborative, and empirical process of cultural mediation can Iranian researchers ensure the conceptual integrity of their adapted instruments and produce cross-culturally valid knowledge.

References

- Acquadro, C., Conway, K., Hareendran, A., Aaronson, N., & European Regulatory Issues and Quality of Life Assessment (ERIQa) Group (2008). Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical

- trials. *Value in Health*, 11 (3), 509–521. <https://doi.org/10.1111/j.1524-4733.2007.00292.x>
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25 (24), 3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>
- Behr, D. (2018). Translating questionnaires for cross-national surveys: A description of a genre and its particularities based on the ISO 17100 categorization of translator competences. *Translation & Interpreting*, 10 (2), 5–20.
- Behr, D., & Braun, M. (2023). How does back translation fare against team translation? An experimental case study in the language combination English–German. *Journal of Survey Statistics and Methodology*, 11 (2), 285–315. <https://doi.org/10.1093/jssam/smac005>
- Behr, D., Braun, M., & Dorer, B. (2016). *Measurement Instruments in International Surveys*. GESIS Survey Guidelines.
- Bowden, A., & Fox-Rushby, J. A. (2003). A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South America. *Social Science & Medicine* (1982), 57 (7), 1289–1306. [https://doi.org/10.1016/s0277-9536\(02\)00503-8](https://doi.org/10.1016/s0277-9536(02)00503-8)
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1(3), 185–216.
- Chen, H. Y., & Boore, J. R. (2010). Translation and back-translation in qualitative nursing research: methodological review. *Journal of Clinical Nursing*, 19 (1–2), 234–239. <https://doi.org/10.1111/j.1365-2702.2009.02896.x>
- Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology*, 46 (12), 1417–1432.
- Harkness, J. A. (2003). Questionnaire translation. In J. A. Harkness, F. J. R. Van de Vijver, & P. Ph. Mohler (Eds.), *Cross-cultural Survey Methods* (pp. 35–56). John Wiley & Sons.
- Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L., Mohler, P. P., Pennell, B.-E., & Smith, T. W. (Eds.). (2010). *Survey methods in multinational, multiregional, and multicultural contexts*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9780470609927>
- Harkness, J. A., Van de Vijver, F. J. R., & Mohler, P. Ph. (Eds.). (2003). *Cross-cultural survey methods*. John Wiley & Sons.
- Heggestad, E. D., Scheaf, D. J., Banks, G. C., Monroe Hausfeld, M., Tonidandel, S., & Williams, E. B. (2019). Scale adaptation in organizational science research: A review and best-practice recommendations. *Journal of Management*, 45 (6), 2596–2627. <https://doi.org/10.1177/0149206319850280>

- International Test Commission. (2017). *The ITC Guidelines for Translating and Adapting Tests (Second edition)*.
https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf
- Jones, P. S., Lee, J. W., Phillips, L. R., Zhang, X. E., & Jaceldo, K. B. (2001). An adaptation of Brislin's translation model for cross-cultural research. *Nursing Research, 50* (5), 300–304 <https://doi.org/10.1097/00006199-200109000-00008>
- Maneesriwongul, W., & Dixon, J. K. (2004). Instrument translation process: a methods review. *Journal of advanced nursing, 48*(2), 175–186.
<https://doi.org/10.1111/j.1365-2648.2004.03185.x>
- McKenna, S. P., & Doward, L. C. (2005). The translation and cultural adaptation of patient-reported outcome measures. *Value in health: the journal of the International Society for Pharmaco-economics and Outcomes Research, 8* (2), 89–91.
<https://doi.org/10.1111/j.1524-4733.2005.08203.x>
- Newmark, P. (1988). *A Textbook of Translation*. Prentice Hall,
- Repke, L., & Dorer, B. (2021). Translate Wisely! An Evaluation of Close and Adaptive Translation Procedures in an Experiment Involving Questionnaire Translation. *International Journal of Sociology, 51* (2), 135–162.
<https://doi.org/10.1080/00207659.2020.1856541>
- Sadeghzadeh, M., Nouhi Jadesi, N. Lost in documentation: professional norms and the gaps in survey translation transparency. *Research Integrity and Peer Review 11, 8* (2026).
<https://doi.org/10.1186/s41073-026-00192-4>
- Sousa, V. D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *Journal of evaluation in clinical practice, 17* (2), 268–274.
<https://doi.org/10.1111/j.1365-2753.2010.01434.x>
- Valdez, D., Montenegro, M. S., Crawford, B. L., Turner, R. C., Lo, W. J., & Jozkowski, K. N. (2021). Translation frameworks and questionnaire design approaches as a component of health research and practice: A discussion and taxonomy of popular translation frameworks and questionnaire design approaches. *Social science & medicine (1982), 278*, 113931.
<https://doi.org/10.1016/j.socscimed.2021.113931>
- van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology / Revue Européenne de Psychologie Appliquée, 54* (2), 119–135. <https://doi.org/10.1016/j.erap.2003.12.004>
- Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). 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, 8* (2), 94–104.
- World Health Organization. (2016). *Process of translation and adaptation of instruments*. WHO.