

Verifying the accuracy of programmed electronic patient-reported outcome (PRO) measures is vital to protect data quality



Lauren Henning, Sloan Rucker, Jahshaun Jenkins, Michelle Carty
QualityMetric, Johnston, Rhode Island, United States

Presented at the 30th Annual ISOQOL Conference, October 18-21 in Alberta, Canada

BACKGROUND

- Electronic administration of PROs has many advantages over paper: less administrative burden, fewer data entry errors, easier execution of skip patterns, safeguards against missing data, and patient preference¹⁻³
- Extensive evidence has shown that electronic PROs (ePROs) are equivalent to their original paper versions if migrated faithfully^{1,4}
- Mistakes made when programming ePROs can produce unreliable data that may result in erroneous findings
 - The multi-national, multi-PRO measure nature of many research studies elevates the risk of mistakes
- To maintain the integrity of data collected via ePRO, many developers require a review process to ensure fidelity to the licensed measure
 - A typical ePRO review process involves checking the programmed survey against its source for accuracy in content and presentation
- Changes in a PRO – whether purposeful or inadvertent – can affect its measurement properties,^{2,5} resulting in compromised data quality

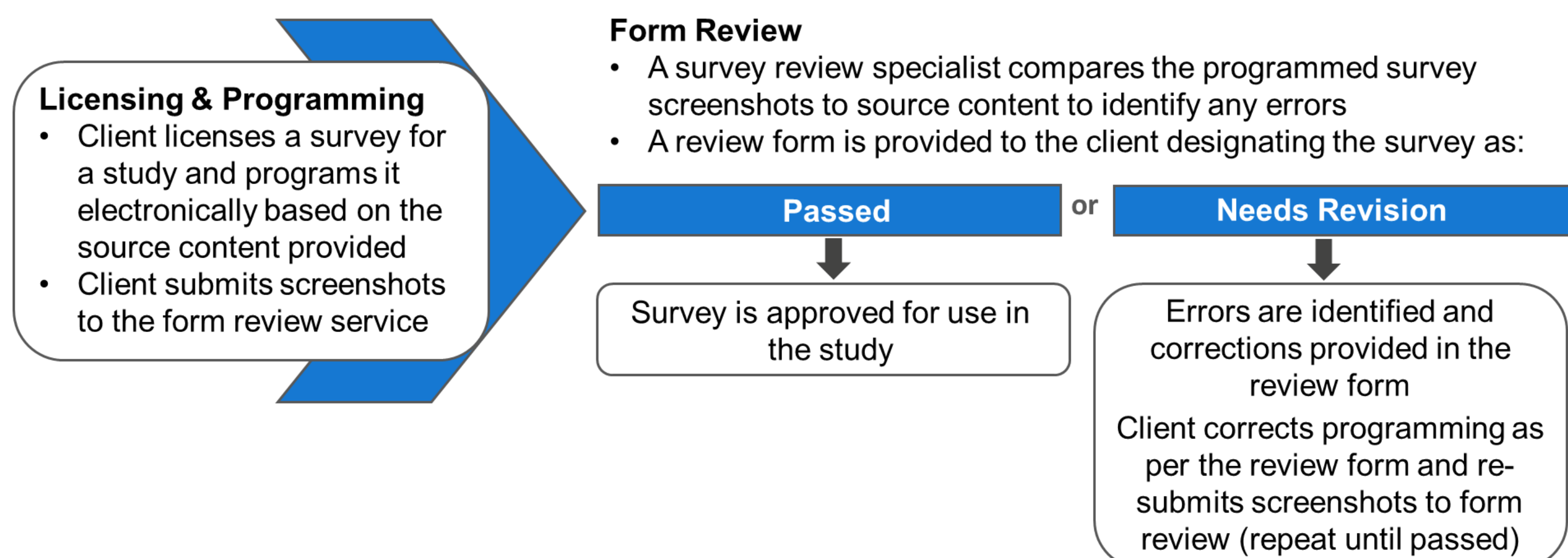
AIM

Evaluate the importance of a review process for ePROs by examining errors made when programming the SF-36v2[®] (SF-36) and SF-12v2[®] (SF-12) Health Surveys

SF-36v2[®] and SF-12v2[®] Health Surveys

- The SF-36 and SF-12 are among the most widely used generic health-related quality of life PROs in clinical trials worldwide
- The SF-36 (36 items) and SF-12 (12 items) measure 8 domains of physical and mental functioning and well-being; scores on all domains are used to calculate summary measures for overall physical and mental health
- The SF-36 and SF-12 are available in 200+ languages and have multiple modes of administration (e.g., paper, electronic) and recall periods
- Electronic formats include tablet, handheld (i.e., mobile device), and ePaper (i.e., paper version online)
- Evidence has shown that electronic versions of the SF-36 and SF-12 are equivalent to paper versions across a variety of formats³
- The review process for the SF-36 and SF-12 (termed “form review”) is depicted in Figure 1

Figure 1. SF-36 and SF-12 Form Review Process



METHODS

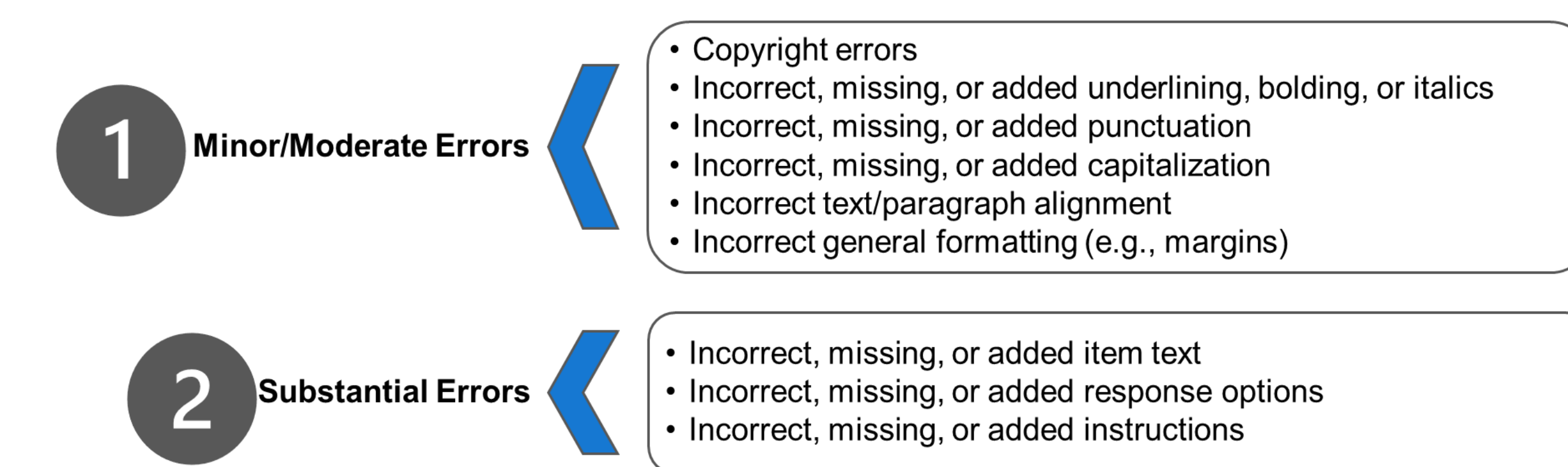
Study Design

- The dataset included initial submissions of all electronic (tablet, handheld, or ePaper), SF-36 and SF-12, acute (1-week recall) and standard (4-week recall) surveys for which form review was completed between January 1 and June 30, 2021
- Each submission review form was coded by 2 independent raters as (0) passed with no errors; (1) minor/moderate error(s); or (2) substantial error using the criteria outlined in Figure 2
 - Errors were classified as substantial if the error could influence the way a respondent answers or interprets survey content²
 - Punctuation and capitalization errors were classified as minor/moderate
 - If a submission contained at least one substantial error, it was coded as (2) substantial even if minor/moderate error(s) were also present
- Coding of each submission review form was checked for agreement amongst raters; discrepancies were resolved in a consensus meeting

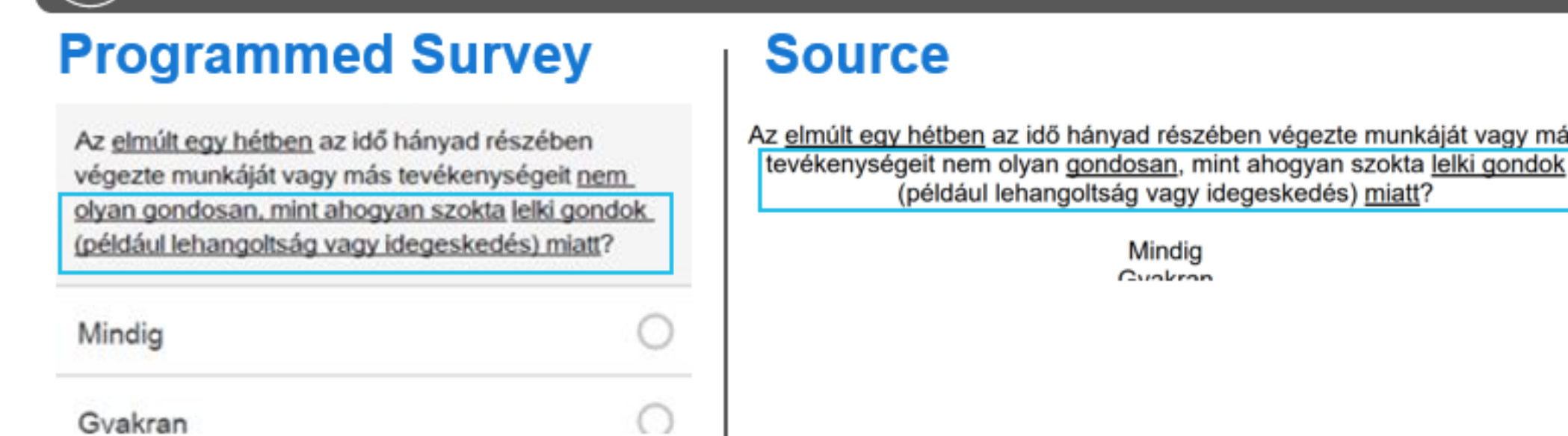
Statistical Analyses

- Descriptive characteristics were summarized for the overall dataset and by survey type (i.e., SF-36, SF-12), language (i.e., English, non-English), and format (i.e., ePaper, tablet, handheld)
- Two-proportion z-tests were used to evaluate differences in the proportion of submitted form reviews with at least one substantial error within survey type, language, and format

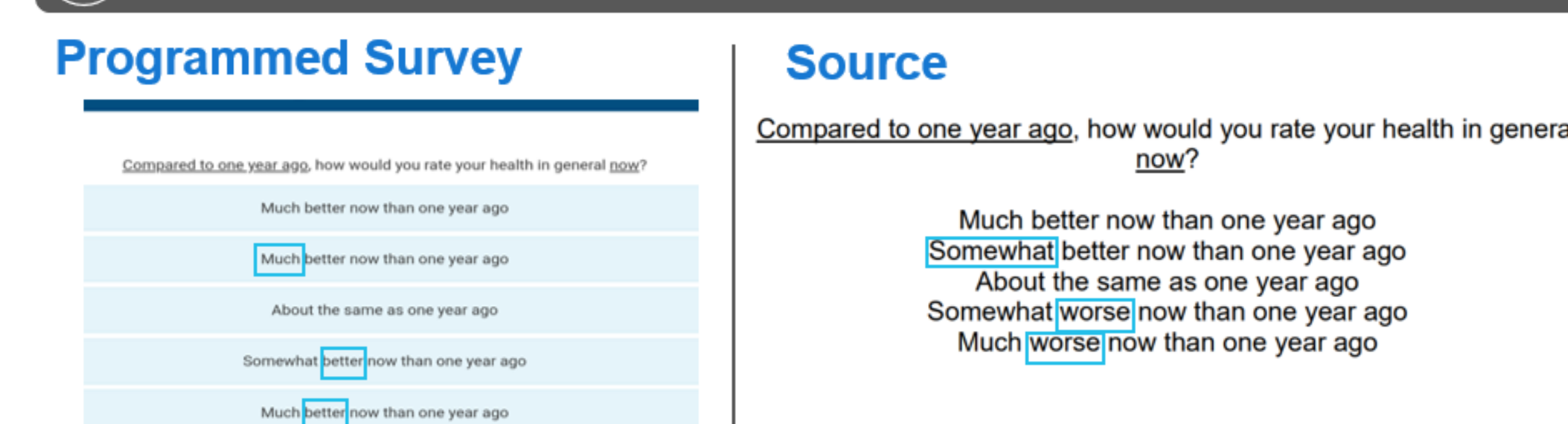
Figure 2. Criteria for (1) Minor/Moderate and (2) Substantial Errors



1 Minor/Moderate Example



2 Substantial Example



RESULTS

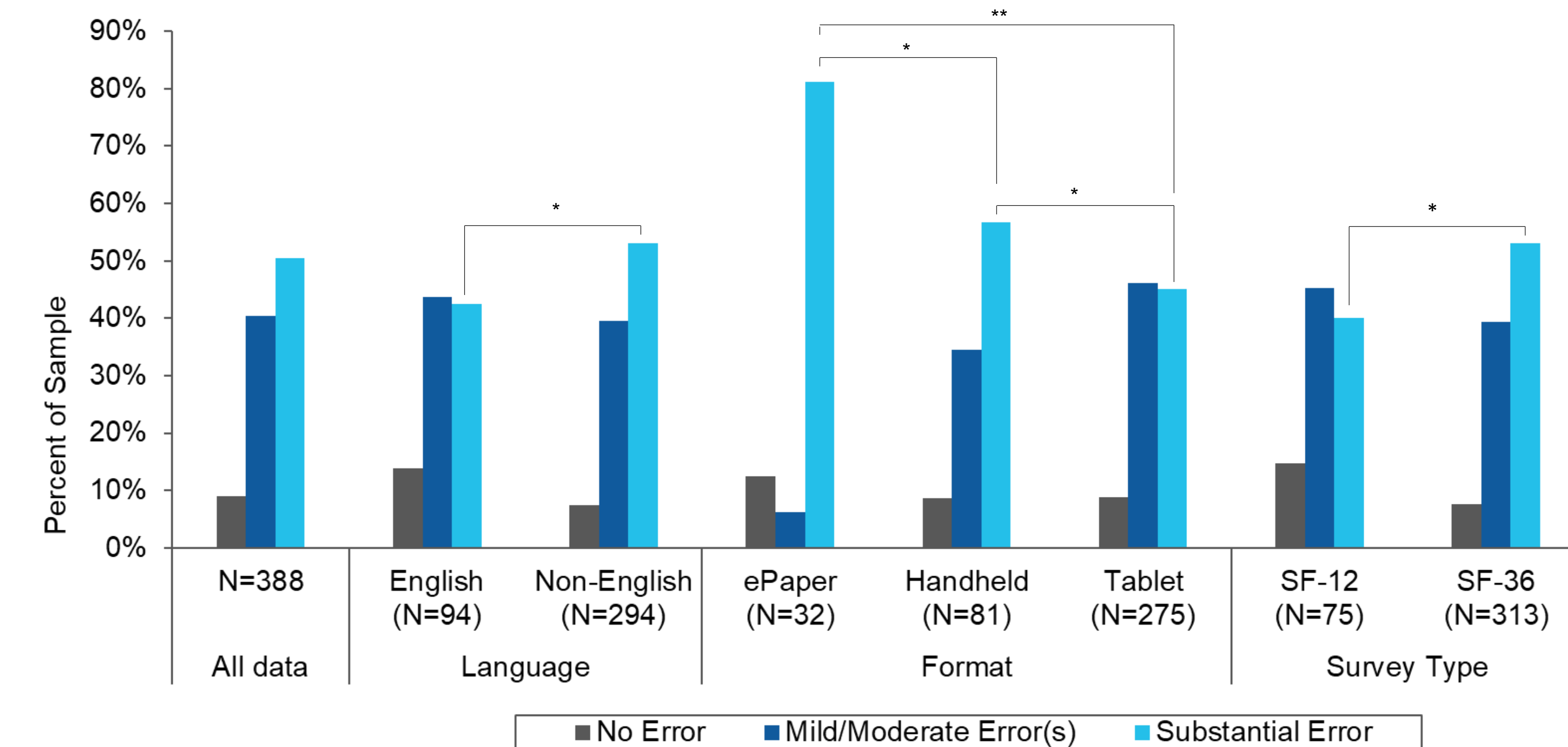
Dataset Characteristics

- Raters coded 388 submission review forms, with 88.1% agreement
 - Only 35 reviews (9.0%) passed on initial submission (i.e., no errors) (Figure 3)
 - 196 reviews (50.5%) identified at least one substantial error and 157 (40.5%) at least one minor/moderate error
- A median of 2 submissions were needed to pass (range: 1-5)
 - 23.0% of surveys needed 3 or more submissions to pass

Group Comparisons

- A significantly larger proportion of reviews identified at least one substantial error for the SF-36 when compared to the SF-12 (p=0.02); non-English language compared to English (p=0.04); ePaper compared to handheld (p<0.01) and tablet (p<.001); and handheld compared to tablet (p=0.03) (Figure 3)

Figure 3. Dataset Characteristics and Within-Group Comparisons



–Asterisks denote a significant difference in the proportion of reviews with at least one substantial error: *p <0.05; **p <0.001

CONCLUSIONS

- Over a 6-month period, 91.0% of electronic SF-36 and SF-12 surveys submitted for review contained errors that required resubmission
 - One-half of those errors were substantial, with the potential to change how a respondent interprets survey content, potentially resulting in unreliable data if form review was not performed
 - A significantly larger proportion of reviews identified at least one substantial error when comparing within survey type, format, and language, suggesting that multiple factors contribute to the likelihood of making a programming error
 - To protect the quality of data obtained from ePRO measures, careful review prior to data collection is essential
- ### Limitation
- The number of errors made during programming the SF-36 and SF-12 was likely underestimated; surveys rejected upon submission to form review as non-reviewable (i.e., programmed survey does not match license) are not tracked and surveys may have gone through programmer quality checks and errors corrected prior to form review

Implications

- A vulnerability in data collection, such as administration of an ePRO that produces unreliable/unusable data, can prove extremely costly in a large clinical trial; a review process can protect against this possibility
- One way to circumvent the need for review of every ePRO, is for PRO owners to partner with ePRO providers to maintain a library of surveys that can be re-used across all studies⁶