

Manual - Checking Tables - Best Practices

Before sending a data tables "book" to a client, it must be checked against the topline and the questionnaire to confirm accuracy. This process uses three documents in combination:

- **The Questionnaire** — the final hardcopy survey, showing all questions, answer codes, skip logic, and flow.
- **The Topline** — the raw frequency output from Survox, cross-tabbed across all banner points.
- **The Tables Book** — the client deliverable, one table per question, with all banner points applied.

These three documents must agree with one another. The checks below confirm they do.

Step 1: Verify the Total N

The first check is simple but critical — confirm that the tables book and the topline are both working from the same dataset.

- In the **tables book**, locate the **Total** column on any table. The value in the top row (total respondents) is your N.
- In the **topline**, find the **Total** column in TABLE 001. Confirm the N matches.

If the two numbers do not match, one of the files is outdated. **The file with the lower N is typically the older version.** Do not proceed until both files reflect the same N.

Step 2: Verify the Banner

The banner — the set of cross-tab columns applied across all tables in the book — is defined once and applied globally. This means it only needs to be verified once, not for every question.

To check the banner:

- Find a question in the tables book that meets all of the following criteria:
 - Asked of **all respondents** (no skip, no filter, no split)

- **Single-response** — multiple-response questions can inflate counts and make the math appear incorrect
- Simple and straightforward — a basic demographic like gender works well
- In the **Total row** of that table, go banner point by banner point across the columns.
- For each banner point, locate the corresponding question in the **topline** and confirm the value matches.

If all banner point values match the topline, the banner is correctly defined and does not need to be re-checked on any other table. If any value is off, something is miscounting in the banner specification and it must be resolved before the book can be delivered.

Step 3: Check Each Table Against the Topline and Questionnaire

With the N and banner both confirmed, work through the tables book question by question. For each table, you are checking two things:

3a. Total Column Matches the Topline

For each answer code row in a table, the value in the **Total column** must match the corresponding value in the topline for that question. If it does not, the table is pulling data incorrectly.

3b. Base N Reflects Skip Logic and Splits

Not every question is asked of every respondent. The questionnaire defines which respondents reach each question — through skips, filters, and routing logic. The table's base N must reflect this.

To verify, find the driving question in the questionnaire — the one whose answer routes respondents into or past the question being checked. Locate that driving question in the topline and identify how many respondents gave the qualifying answer. That count should match the base N of the dependent table.

Example: If the questionnaire says "*If Q2 = 3, continue to Q3; otherwise skip to Q4*", and the topline shows 92 respondents answered 3 on Q2 — then the base N on the Q3 table should be 92.

Split-assigned questions follow the same principle. A split is not driven by another question — it is a pre-assigned quota portion of the total N. A two-way split (A/B) is approximately 50%/50%. A three-way split (A/B/C) is approximately 33%/33%/33%, and so on. Some studies use multiple independent split sets, where a respondent may belong to one split from each set (e.g., they are in Split A of one set and Split D of another). In all cases, the base N for a split-assigned question should match the corresponding split count shown in the topline.

If the base is wrong, either the skip logic or split assignment was programmed incorrectly in the tables.

Things to Watch For

Multiple-Response Questions

Some questions allow respondents to select more than one answer. When checking these tables, the sum of all row values will often **exceed** the total N — and that is expected, not an error. This happens because the Total row counts **respondents**, while each answer code row counts **responses**. A single respondent who selected three answers contributes once to the total N but three times across the individual rows.

As long as the sum of the row values is **equal to or greater than** the total N, the table is correct. The same logic applies to percentages — individual row percentages will not add up to 100% on a multiple-response question, and they are not supposed to.

Note: The hardcopy questionnaire does not always clearly label whether a question is single or multiple response. When in doubt, check the original programming specifications or confirm with the programmer.

Mean, Median, and Other Calculated Statistics

Means, medians, and similar statistics are calculated automatically by the system. These do not need to be manually verified as part of the standard table check.

Summary Tables

Some tables book deliverables include summary tables — a single table that consolidates results from multiple related questions. A common example is a battery of yes/no brand awareness questions summarized into a single "yes" table, rather than delivering each question individually.

When checking a summary table, each row should match the corresponding individual question's value in the topline. If the individual questions are also present in the tables book and have already been checked, those verified values can be used as the reference instead of going back to the topline.

If the individual questions were not included in the client deliverable, the topline is the only reference — apply the same row-by-row total column check used for all other tables.

Summary Banner Points

Some banner points are themselves summaries — combining mentions of a specific code across multiple questions. For example, a "User of Brand X" banner point might count any respondent who mentioned Brand X in Q1, Q3, Q8, Q11, Q15, or Q26.

Because a respondent only counts once regardless of how many of those questions they mentioned Brand X in, this behaves like a unique respondent count — not a response count. The total for a summary banner point will land somewhere between the highest single-question value and the sum of all contributing question values, depending on how much respondent overlap exists between questions.

Manual verification of the exact number is not practical without knowing the overlap. However, the following rule always applies:

The summary banner point total cannot be lower than the highest value seen in any single contributing question.

Example: If Q3 had 40 respondents mention Brand X, and Q7 had 6 — the summary banner point must show at least 40. If 3 of those 6 in Q7 were also among the 40 in Q3, the correct total is 43 unique respondents. It cannot be 30, and it cannot exceed 46 (the raw sum).

If the summary banner point value falls below the highest individual question value, it is wrong and must be corrected before delivery.

Weighted Data

Some studies use weighted data. This is typically requested by clients who want to collect data as cost-effectively as possible during fielding, but want the final deliverables to reflect the natural demographic composition of the surveyed area — usually based on census data for that region. Rather than controlling for age, gender, race, or other demographics during data collection, the DP team assigns a relative weight value to each case after fielding. When the final tables are run, Survox applies those weight values as multipliers across the data.

Weighted tables are inherently difficult to check manually. Any time whole numbers are multiplied by percentage-based weight values, rounding comes into play at the individual case level — and those rounding differences accumulate across the dataset. A weight of 1.1 rounds down to 1, and a weight of 2.4 rounds down to 2, but their combined value of 3.5 rounds up to 4. Small discrepancies in weighted totals are expected and are not necessarily errors.

For this reason, **unweighted tables should always be checked first.** Unweighted data is a true 1:1 ratio of responses to respondents, which makes discrepancies easier to identify and confirm. Once the unweighted tables are verified clean, weighted tables can be reviewed with the understanding that minor rounding variance is normal.

Statistical Testing

When requested by the client, statistical significance testing is applied to the tables. MAXimum Research uses All-Pairs testing at the 90% and 95% confidence levels, testing all columns against one another simultaneously.

In the tables, columns are labeled alphabetically — the Total column is typically labeled **A**, and each subsequent banner point column is labeled **B**, **C**, **D**, and so on. Where a statistically significant difference exists between columns, the letter of the column being outperformed appears in the cell of the higher-performing column.

Stat testing annotations are generated automatically by the system. There is no standard manual method to verify whether a specific letter flag is correct. When clients ask about stat testing, the annotations can be pointed to and explained at a high level, but no further interpretation or validation of individual flags is expected as part of the table checking process.

Completing the Check

Once all tables have been verified, notify the programmer by email that the check is complete and the book is clear for client delivery — or flag any discrepancies found for correction.

Revision #4
Created 2026-03-27 19:49:17 UTC by Admin
Updated 2026-04-13 20:31:05 UTC by Admin