Turas MaxDiff

Transparent. Trustworthy. Tangible insights.

Reveals true preference rankings—which features, messages, or attributes matter most to your customers. MaxDiff gets clean preference data through simple "best" and "worst" choices, avoiding the scale-use bias that plagues rating questions.

What This Module Does

Traditional ranking questions are hard when you have many items (10+ options). Respondents struggle to rank 15 features from most to least important, and direct ratings suffer from "everything is important" bias. MaxDiff solves this with repeated best/worst choices.

Instead of ranking 15 features, respondents see sets of 4-5 items at a time and pick the MOST and LEAST important. Repeating this 10-12 times across different combinations, the module calculates definitive preference scores with measurable gaps between items.

What You Get

Aggregate Outputs

- **Preference scores:** 0-100 scale for easy interpretation and comparison
- **Definitive rankings:** Clear order from most to least preferred
- **Statistical significance:** Which items are significantly different from each other
- **Probability scores:** Chance each item is chosen as best

Individual Outputs (HB Method)

- **Personal preference scores:** Each respondent's priorities—enables segmentation
- **Individual rankings:** How each person differs from the population average
- **Credible intervals:** Uncertainty around individual estimates

How It Works

Best/Worst Method

From a set of 4-5 items, pick the MOST and LEAST important. This forces real trade-offs (can't say everything is "9 out of 10"), is cognitively easier than ranking, and provides 2 data points per task.

Conditional Logit (Aggregate)

Models choice probabilities based on item utilities. Uses survival::clogit (Mayo Clinic's implementation with 30+ years of clinical validation). Fast, robust, provides population-level preferences.

Hierarchical Bayes (Individual-Level)

Optional Stan-based estimation provides individual utilities. Uses HMC sampling with full convergence diagnostics (Rhat, ESS, divergences). Enables sophisticated segmentation by preference patterns.

Packages Used

All packages are peer-reviewed, open-source R packages available on CRAN.

Package	Why We Use It
survival::clogit	Mayo Clinic's conditional logit—30+ years clinical validation
cmdstanr	Andrew Gelman's Stan team for HB estimation (optional)
AlgDesign	Optimal experimental design for balanced item sets

Strengths

- ✓ **More discriminating:** Better at separating preferences than rating scales
- ✓ **Less fatigue:** Easier for respondents than ranking 15+ items
- ✓ **Individual-level (HB):** Bayesian method provides personalized preference scores
- ✓ **Avoids scale bias:** Doesn't suffer from "everything is important" problem
- ✓ **Large item sets:** Can handle 20-30 items efficiently
- ✓ **Flexible installation:** Works with or without Stan (empirical Bayes fallback)

Limitations

- ⚠ **Survey length:** Needs 8-15 choice tasks per respondent for stable estimates
- ⚠ **Requires careful design:** Task sets must be balanced and orthogonal
- ⚠ **Assumes unidimensional:** All items rated on same underlying dimension
- ⚠ **Stan computation time:** Full HB estimation requires setup time and computational resources

Sample Size Requirements

Aggregate (clogit): minimum n≥150, recommended n≥250. Individual-level (HB): minimum n≥200, recommended n≥300. Need 8-15 tasks per respondent with 4-5 items shown per task.

Best Use Cases

Ideal For

- ✓ Feature prioritization (which features to build first)
- ✓ Message testing (which claims resonate most)
- ✓ Brand attribute importance
- ✓ Benefit prioritization for positioning
- ✓ Any "what matters most" question with 10+ options

Not Ideal For

- ⚠ Price-feature trade-offs (use Conjoint instead)
- ⚠ When you need only 5-6 items ranked (simple ranking may suffice)
- ⚠ When items span multiple dimensions (importance vs. satisfaction)

Key Takeaways

- ✓ MaxDiff produces cleaner rankings than rating scales—forces real discrimination
- ✓ Preference scores (0-100) show relative importance with measurable gaps
- ✓ Statistical significance testing shows which items truly differ
- ✓ HB enables segmentation by preference patterns—different things matter to different people
- ✓ Large item sets (15-25) work well; very large (30+) increases respondent burden

The Bottom Line

MaxDiff is the gold standard for preference measurement when you have many items to evaluate. It produces cleaner, more discriminating results than rating scales while being easier for respondents than ranking tasks. The conditional logit approach provides rigorous statistical estimation, while hierarchical Bayes enables individual-level insights for sophisticated segmentation. Transform "everything is important" into "here's your definitive priority ranking."

About The Research LampPost

Turas is developed and delivered by **The Research LampPost**, an organisational member of the Southern African Marketing Research Association (SAMRA). Duncan Brett is a SAMRA accredited researcher with 30+ years of market research experience. If you have questions about whether Turas is right for your project, if there is interest in the statistical documentation or module-specific information, or you would like a copy of our credentials, please get in touch at duncan@researchlamppost.co.za

Ready to discuss your project?

duncan@researchlamppost.co.za