Designing Reliable Lanham Act Surveys: Eveready & Squirt — A Short Primer for Attorneys

By Jonathan Z. Zhang, Ph.D. | Dr. Ajay Menon Professor of Marketing, Colorado State University

Email: jonathan.zhang@colostate.edu | Phone: 646.750.2773

This primer outlines when to use Eveready vs. Squirt designs, how to select the right universe, what courts typically expect on controls and stimuli realism, and how to report results clearly for admissibility under Rule 702. It is written for litigators who need a dependable, transparent survey process with quick turnaround.

1) When to use Eveready vs. Squirt (lineup)

- Use Eveready when the senior mark is well known to the target consumers and the challenged product is encountered without side-by-side comparison. Respondents see the junior product and are asked source/affiliation questions without being shown the senior mark.
- Use Squirt/lineup when consumers are likely to encounter the marks together (e.g., same shelf, search results, catalogs) or where the senior mark is not sufficiently famous to support unaided recognition. Respondents view both stimuli in a realistic lineup, then answer source/affiliation questions.
- Rule of thumb: pick the design that best mirrors the real-world exposure and decision context you can substantiate with marketplace evidence (shelf photos, search pages, ad placements, site analytics).

2) Defining the survey universe (who to sample)

- Sample only those prospective purchasers or influencers reasonably likely to encounter and consider the product/service at issue.
- Document the inclusion logic: category usage, purchase frequency, brand/category familiarity, geography, age/role constraints.
- Screen out professionals/uninvolved respondents; use incidence checks; record drop rates so the court can see how you arrived at the final sample.

3) Stimuli realism and presentation

- Use market-accurate stimuli (current packaging, labeling, disclaimers, site pages). Avoid ‘clean room’ images unless that is how buyers actually see the product.

- Match the exposure channel (in-person shelf photos, mobile web screenshots, sponsored search pages, social feeds) and device form factor when feasible.
- Randomize order, placement, and any filler items; keep exposure time reasonable and consistent with real shopping.

4) Question design (source, affiliation, and confusion)

- Lead with open-ended source/affiliation/sponsorship questions before any brands are named (e.g., "Who do you think makes this?" "Do you think it is affiliated with or approved by any company? Which?").
- Follow with closed-ended checks to assess confusion, mistaken affiliation, or approval. Avoid leading language (no brand lists, no hints).
- Capture verbatim responses and code to pre-registered categories; keep coders blind to cell assignment.

5) Controls & noise adjustment

- Include a control cell that removes or alters the allegedly infringing features while holding everything else constant (e.g., an innocuous brand/packaging or a modified mark).
- Report net confusion = (test confusion − control confusion) to address background noise and guessing.
- Pre-specify control logic in the protocol; justify it with branding principles (similarity, distinctiveness, dominant features).

6) Sample size, power, and quality checks

- Typical target: n≈200–300 per cell, adjusted for incidence and effect size; justify with a simple power calculation.
- Quality screens: attention checks, straight-lining flags, time thresholds, open-ended coherence checks, duplicate IP/device filters.
- Document panel sourcing, quotas, and completion rates; maintain a fieldwork log and chain-of-custody for stimuli files.

7) Analysis & reporting

- Compute and disclose gross and net confusion with 95% confidence intervals; include subgroup analyses only if pre-specified or clearly labeled as exploratory.
- Use standard tests (z-test for proportions, chi-square, or logistic regression) and report exact p-values.
- Present a clear table: sample sizes, confusion counts, net confusion, CI, and control description. Include representative coded verbatims.

8) Common pitfalls (and how we prevent them)

- Unrealistic stimuli or exposures → Use verified shelf/search evidence and device-specific screenshots.

- Leading questions → Keep opens first, neutrally worded, and pre-test for comprehension.
- Bad controls → Justify controls in writing and pilot them to ensure they don't introduce new confusion.
- Poor coding → Double-code a sample, reconcile disagreements, and keep coders blind to condition.

9) Deliverables & timeline (typical)

- Protocol & annotated questionnaire (7–10 days after record review).
- Fieldwork & QC memo (5–7 days).
- Expert report with methods, results, exhibits, and appendices (10–14 days post-field).
- Production set: instrument, stimuli, field logs, de-identified dataset, codebook, screenshots/photos.

10) Quick checklist for admissibility (Rule 702)

- Methods tied to facts of the case; survey universe reasonably matches likely purchasers.
- Reliable principles and methods (accepted designs; documented controls; transparent coding).
- Reliable application (pre-registered protocol; QC; reproducible analysis; full work-product preserved).