

AI SHOPPING DECISION AUDIT

AI Shopping Decision Audit

P1-01

P1-02 Smart Birdy - Deluxe Grey Gold | United States | Use: proposal | Researched 23 Aug 2026

THE 30-DAY RESOURCE DECISION

P1-03 For the next 30 days, fund one exact-set packing demonstration for beauty, fashion, and event trips; validate the business-trip one-bag promise with a physical pack-and-sizer test; spend \$0 promoting the two-case, one-hand scene.

FUND

Exact-set proof

P1-04 FUND - Beauty / fashion / event trips. Product structure is the strongest fit: the suitcase supports shoes, clothing, and a cooled styling-tool loop; the \$1,499 Luggage Set adds jewelry and toiletry accessories. Smart Birdy appeared in 0/4 sampled answers while BÉIS Carry-On Roller appeared in 4/4. Fund accurate, independent evidence - not the absence count by itself.

VALIDATE

One-bag test

P1-05 VALIDATE - 2-4 day business trips. Laptop, passport, and removable-charging features are relevant, but external quick access is unconfirmed, the battery is Set-only, and 19 x 16 x 9 in conflicts with some US airline limits. Prove the workflow before promotion.

AVOID

Two-case stop

P1-06 AVOID - Moving two packed rollers with one hand. Smart Birdy's stacking sleeve is not the same as a positive mechanical lock; the current Dual Stack price is \$2,000 and fulfillment is unclear. TACH appeared in 4/4 and BÉIS in 2/4 with clearer connection mechanisms and lower prices.

P1-07 Scope: one brand, one hero product, the US market, three buying scenes, and 12 independent model runs (ChatGPT x2 + Gemini Flash x2 per scene). Research is independent public-data analysis and was not commissioned, reviewed, or endorsed by Smart Birdy.

P1-08 No ranking, citation, recommendation, traffic, or sales outcome is promised. Model inclusion counts are dated observations, not market share, win rates, or stable recommendation probabilities.

MEETING ANSWER

One scene to fund. One to prove. One to leave alone.

AI SHOPPING DECISION AUDIT

Fund / Validate / Avoid

P2-01

P2-02 Each decision combines current product facts, independent evidence, model observations, and the strongest counterevidence. No composite visibility score is used.

FUND**S2 | Beauty / fashion / event**

P2-03 BUYER JOB - A US woman on a 2-4 day beauty, fashion, or event trip wants one polished carry-on that separates shoes, a hair tool, jewelry, toiletries, and clothing to reduce extra bags.

P2-04 CURRENT SET - Smart Birdy 0/4; BÉIS Carry-On Roller 4/4; all other candidates no more than 2/4. FIT - The suitcase has a shoe compartment, hot-tool loop, and clothing storage; jewelry and toiletry accessories belong to the \$1,499 Set, not the \$999 suitcase.

P2-05 COUNTEREVIDENCE - Smart Birdy does not claim heat resistance; empty weight is 8 lb; current independent durability evidence remains thin. WHY FUND - This is the clearest factual-fit / candidate-absence mismatch and can be tested with one bounded evidence asset.

VALIDATE**S1 | Business-trip one-bag**

P2-06 BUYER JOB - A US woman on a 2-4 day business trip wants a polished hard-sided carry-on with quick laptop and passport access plus removable charging, without a separate personal bag.

P2-07 CURRENT SET - Smart Birdy 0/4; July Carry On Pro family 3/4; July Capsule family 3/4; Travelpro 3/4. FIT - Laptop and passport storage exist, but external quick access is not established; the battery is Set-only.

P2-08 COUNTEREVIDENCE - 19 x 16 x 9 in exceeds American Airlines' 18 x 14 x 8 in personal-item limit and its 14 in carry-on width; it also exceeds Delta's 14 in carry-on width. WHY VALIDATE - Any one of access, size, packed weight, or one-bag behavior can reverse the budget decision.

AVOID**S3 | Two packed rollers**

P2-09 BUYER JOB - One US traveler regularly moves two packed rolling cases alone and wants a secured stack or physical connection that can be controlled with one hand.

P2-10 CURRENT SET - TACH 4/4; BÉIS 2/4; Smart Birdy Dual Stack 1/4. FIT - Smart Birdy uses the Deluxe trolley sleeve over an Essential base, but this is stacking support rather than a BÉIS/TACH-style positive lock.

P2-11 COUNTEREVIDENCE - There is no current independent full-load stability test; the \$2,000 price and unclear fulfillment compare poorly with clearer, lower-priced systems. WHY AVOID - Better model visibility would not remove the product-mechanism and price disadvantage.

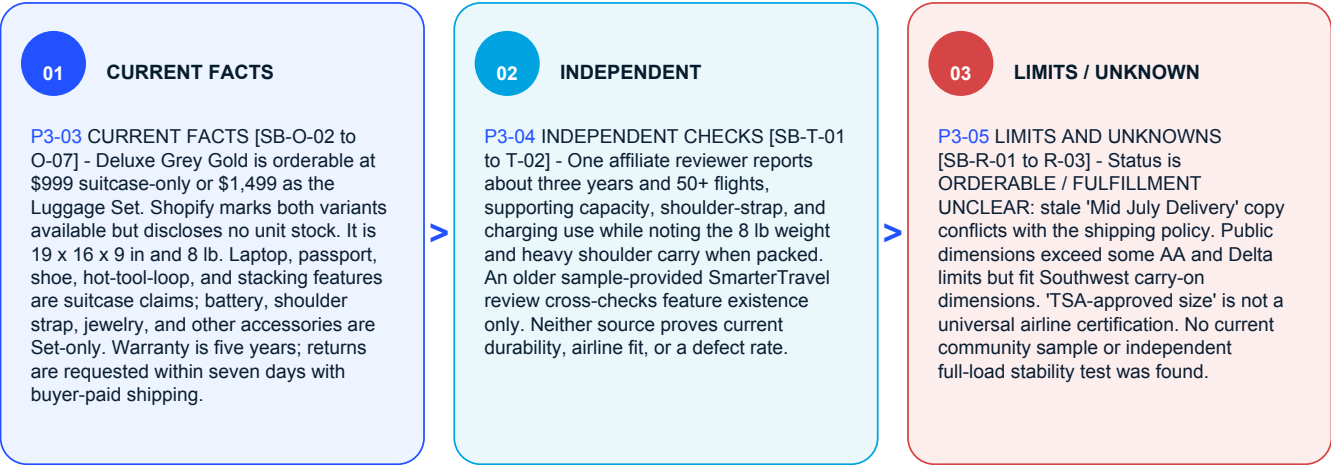
P2-12 Count note: x/4 means the named candidate appeared in x of four answers for that scene on 23 Aug 2026. Multiple products from one brand may appear in one answer. Counts do not establish superiority.

AI SHOPPING DECISION AUDIT

Evidence Chain

P3-01

P3-02 Decision logic: current facts -> independent checks -> limits and unknowns -> 12 dated model receipts -> resource choice.



12-RUN MODEL EVIDENCE RECEIPT

P3-06 MODEL OBSERVATIONS [12 runs] - Smart Birdy appeared S1 0/4, S2 0/4, S3 1/4. The static receipt table records every run, provider, sampling window, target inclusion, repeated candidates, and strongest limitation. Counts can be recomputed without a provider account; full outputs remain in the internal delivery ledger. Product facts override model claims.

RUN / PROVIDER	WINDOW	SMART BIRDY	REPEATED CANDIDATES	STRONGEST OBSERVATION
S1-C1 ChatGPT	22:14-16	NO	July; Travelpro	No strict match; July width conflict
S1-C2 ChatGPT	22:19-22	NO	July; Travelpro	No strict match; battery/passport/size gap
S1-G1 Gemini	22:16-18	NO	July; Arlo Skye	Treated about 15 in as 14-in compatible
S1-G2 Gemini	22:22-24	NO	July; Travelpro	Strict-match claim conflicts with cited width
S2-C1 ChatGPT	22:19-22	NO	BEIS	No strict match; added heat-resistance test
S2-C2 ChatGPT	22:19-22	NO	Samsonite; BEIS	No strict match; jewelry/hot-tool gap
S2-G1 Gemini	22:22-24	NO	BEIS	No strict match; conflicting BEIS prices
S2-G2 Gemini	22:22-24	NO	BEIS	No strict match; separate hot-tool sleeve
S3-C1 ChatGPT	22:19-22	NO	BEIS; TACH	Both treated as secured full matches
S3-C2 ChatGPT	22:19-22	YES	TACH; BEIS	Smart Birdy partial; stale \$1,400 price
S3-G1 Gemini	22:22-24	NO	TACH	TACH dominated; Smart Birdy absent
S3-G2 Gemini	22:22-24	NO	TACH	TACH most direct; other systems partial

RESOURCE INFERENCE

P3-07 INFERENCE - S2 is the best current evidence-investment mismatch; S1 has a plausible use case but an unresolved physical promise; S3 has a product and price disadvantage that content cannot repair. This inference changes if the physical tests below fail or new independent evidence reverses the comparison.

P3-08 Evidence labels: CONFIRMED = current direct source; THIRD-PARTY = independent but bounded; CONFLICT / UNKNOWN = unresolved and decision-relevant; MODEL OBSERVATION = sampled output, not product truth; INFERENCE = a reversible resource judgment.

AI SHOPPING DECISION AUDIT

30-Day Action Plan

P4-01

P4-02 Review owner and final decision authority: Brand owner. The agency maintains a dated client action log with source links, test receipts, and the signed Continue / Adjust / Stop decision. All three scenes are reviewed on 22 Sep 2026; no action becomes an ongoing content program before that decision.

FUND

P4-03 FUND DRI - Agency strategist; medium effort. Brand supplies and approves current Deluxe Luggage Set facts. One independent reviewer records the pack test. Agency publishes one crawlable exact-set page and the four-run re-sample; Brand owner signs the decision. Test shoes, a cooled styling tool, jewelry/toiletry accessories, clothing, and suitcase-only / Set boundaries. Never claim heat resistance.

P4-04 PASS EVIDENCE / GATE - The review and specification page must name Deluxe Luggage Set, separate Set-only components, state cooled-tool / no-heat-resistance limits, and show the listed load replacing at least one separate shoe, beauty, or toiletry bag versus the reviewer's normal setup. CONTINUE if all pass and at least 2/4 re-sampled answers identify the exact Set with correct bundle facts. ADJUST if the physical/page tests pass but only 0-1/4 do so or entity/bundle errors remain. STOP if no extra bag is removed or either safety/bundle boundary cannot be published.

VALIDATE

P4-05 VALIDATE DRI - Brand operations lead; low-medium effort. One independent business traveler packs a laptop, passport, Set-only removable charging hardware, and 2-4 days of clothing. Agency records quick-access video, loaded weight, and physical-sizer receipts for American Airlines and Delta; Brand owner signs the decision. No promotion before review.

P4-06 PASS EVIDENCE / GATE - Laptop and passport are reached without opening the main shell; one case holds the full listed load with no separate personal or laptop bag; battery copy says Set-only and removable; the loaded case enters each named sizer fully without force. CONTINUE if all four conditions pass for both American Airlines and Delta. ADJUST if the workflow passes but only a documented narrower carrier such as Southwest passes. STOP if no named airline passes, a separate bag is still needed, or battery/access evidence fails.

AVOID

P4-07 AVOID DRI - Brand owner; low effort; paid promotion \$0. Agency only maintains accurate stack-vs-lock, price, and fulfillment copy; Brand owner signs the 22 Sep decision. CONTINUE AVOID unless an independent tester completes three identical loaded-route trials against TACH V3.1 and BÉIS Hybrid Attache + Carry-On Roller with zero tip/detach events, median one-hand time no slower than either control, and documented orderable price/ship date no worse than at least one control. ADJUST only to post-purchase education; REVERSE to Validate only if every threshold passes.

DO NOT FUND	OVERALL STOP
P4-08 DO NOT FUND THIS CYCLE - Generic 'best carry-on' content; universal personal-item or carry-on claims; 'TSA-approved size' language; heat-resistant claims; durability or quality superlatives; creator / PR spend for the two-case scene; a broad monitoring dashboard; or any claim that a model count predicts sales.	P4-09 Overall stop condition: if the independent S2 packing test cannot demonstrate fewer bags without relying on undocumented capability, downgrade S2 immediately and do not substitute more content, more prompts, or a broader campaign for the failed product proof.

AI SHOPPING DECISION AUDIT

Sources, Method & Stop Conditions

P5-01

DIRECT SOURCES

P5-02 SOURCES - Fifteen direct links support the 23 Aug 2026 evidence package. Official pages support current product, policy, airline-rule, and control-product statements; the two independent reviews are explicitly bounded.

[SB-O-01 Smart Birdy home](#)
[SB-O-02/04 Deluxe Grey Gold product](#)
[SB-O-03 Shopify product endpoint](#)
[SB-O-05 Smart Birdy FAQ](#)
[SB-O-06 Smart Birdy shipping](#)
[SB-O-07 Warranty & returns](#)
[SB-O-08 Smart Birdy Dual Stack Set](#)
[SB-T-01 A Solo Woman Traveling review](#)
[SB-T-02 SmarterTravel review](#)
[SB-R-01 American Airlines carry-on](#)
[SB-R-02 Delta carry-on baggage](#)
[SB-R-03 Southwest carry-on policy](#)
[CTRL-01 BEIS Hybrid Carry-On Roller](#)
[CTRL-02 July Carry On](#)
[CTRL-03 TACH V3.1 Carry-on](#)

METHOD & BOUNDARIES

P5-03 METHOD - Three fixed English prompts; four independent new conversations per scene: ChatGPT x2 (logged out; visible version unavailable) and Gemini Flash x2 (logged in). US / English context; network region uncontrolled. Page 3 preserves 12 static run receipts, including unfavorable results. x/4 is dated answer inclusion, not a score. Current facts override model claims.

P5-04 BOUNDARIES - Prices, orderability, airline rules, and model outputs can change after 23 Aug 2026. Available=true does not prove unit stock or immediate fulfillment. Suitcase-only and Set contents must not be merged. Airline fit depends on route, carrier, aircraft, and enforcement. Stop or downgrade when physical evidence fails; do not compensate with more prompts or content. No ranking, citation, recommendation, traffic, or sales outcome is promised.

EXACT PROMPTS USED IN ALL FOUR RUNS PER SCENE

P5-05 S1 PROMPT - Recommend up to 5 specific rolling luggage products currently available to a US shopper for this buying job: a woman taking 2-4 day business trips wants one polished hard-sided carry-on with quick-access storage for a laptop and passport plus a removable battery or included charging solution. She wants to avoid carrying a separate personal bag. For each recommendation, name the exact product, explain which required features are supported by current public information, state whether the charging hardware is included or sold separately, give overall dimensions and empty weight when verifiable, and flag airline compatibility by specific airline rules rather than calling a product 'TSA-approved.' Do not infer undocumented features. If no product is a strict match, say so and list the closest partial matches with their missing requirements.

P5-06 S2 PROMPT - Recommend up to 5 specific rolling luggage products currently available to a US shopper for this buying job: a woman taking 2-4 day beauty, fashion, or event trips wants one polished carry-on that can separately organize shoes, a hot hair tool, jewelry, toiletries, and clothing so she can reduce extra bags. For each recommendation, name the exact product, explain which required compartments or accessories are supported by current public information, state which items are included with the suitcase versus available only in a bundle or as separate accessories, and give current price, overall dimensions, and empty weight when verifiable. Do not infer undocumented features. If no product is a strict all-in-one match, say so and list the closest partial matches with their missing requirements.

P5-07 S3 PROMPT - Recommend up to 5 specific luggage products or systems currently available to a US shopper for this buying job: one traveler regularly needs to move two packed rolling suitcases alone through airports and wants the cases to stack or physically connect so both can be controlled with one hand. For each recommendation, name the exact product or set, explain the verified connection or stacking method, state which pieces and straps are included versus sold separately, and give current price, carry-on dimensions, and empty weight when verifiable. Distinguish a true built-in or secured connection from merely balancing a bag on a trolley sleeve. Do not infer undocumented features. If no option fully meets the job, say so and list the closest partial matches with their missing requirements.