Can a rep trust the policy lookup?

23 / 24

in-scope policy questions matched to the intended passage once the lookup learned shoppers' words, up from 12 of 24 with direct wording.

+11

net additional correct matches

4 of 6

off-topic prompts referred to a person

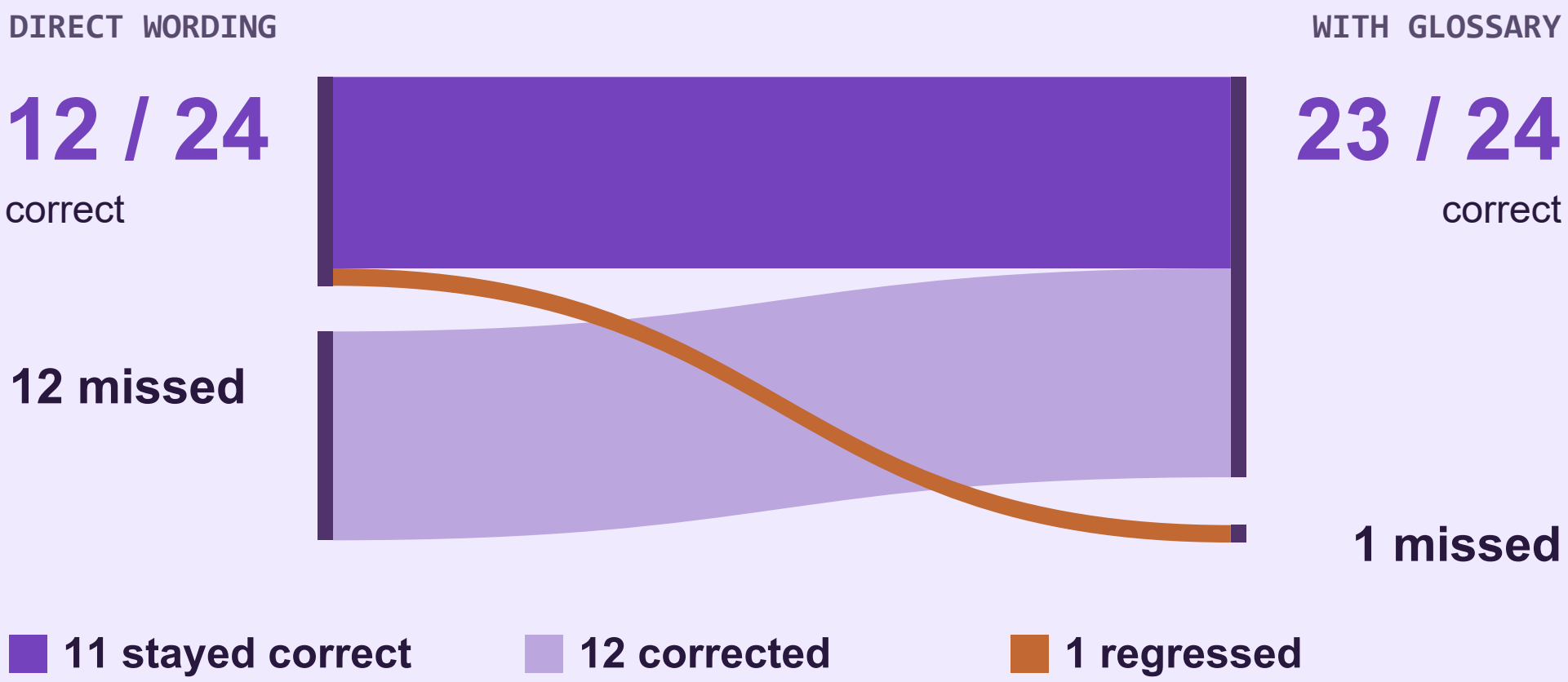
120 h

monthly capacity at 60% adoption, 2 minutes a ticket

Source: Authored support-policy benchmark • 30 prompts across eight policies

Where did the extra matches come from?

The same 24 in-scope questions were scored before and after adding the shopper glossary.



Net change was 11 additional correct matches, from 50.0% to 95.8%.

Source: paired in-scope rows in BENCHMARK.csv. The six off-topic prompts are evaluated separately.

Which policy still breaks?

All eight handbook sections had three questions. Each square preserves the same prompt position across both columns.

EXPECTED SECTION		DIRECT WORDING	WITH GLOSSARY
R-01	Unopened returns	× + ×	+ + +
R-02	Opened returns	+ + ×	× + +
W-01	Damaged deliveries	× + ×	+ + +
B-01	Duplicate charges	× + +	+ + +
D-01	Address changes	× + +	+ + +
S-01	Subscription pauses	× + ×	+ + +
G-01	Gift returns	+ + ×	+ + +
I-01	International duties	× + ×	+ + +

The remaining error sent the opened-item question to R-01, the unopened-return section.

Source: BENCHMARK.csv. + = intended passage; × = wrong passage or referral on an in-scope prompt.

When does the desk answer off topic?

A support lookup must find an intended passage and decline a request that the handbook does not cover.

SIX OFF-TOPIC PROMPTS

R R R R ! !

4 referred

The lead received the intended handoff.

2 answered

Both returned a policy outside the request.

FALSE MATCHES

Sunday opening hours

W-01 / Damaged deliveries

Business ownership

B-01 / Duplicate charges

Off-topic referral accuracy remained 4/6 in both runs. A better passage score alone cannot fix this.

Source: six rows with expected_policy = REFER. The sample is small; replay new off-topic cases before rollout.

How much capacity could adoption release?

Scenario model for 6,000 eligible tickets per month. The cells show capacity in hours, before any service-level change.

ADOPTION	1 min / ticket	2 min / ticket	3 min / ticket
30%	30 h	60 h	90 h
60%	60 h	120 h	180 h
90%	90 h	180 h	270 h

Middle case 6,000 × 60% × 2 minutes ÷ 60 = 120 hours per month

---

Measure eligible volume, time saved and referral accuracy on real tickets before changing staffing.

Source: project planning assumptions. The 1- and 3-minute columns vary the base 2-minute assumption.