



— BENCHMARK REPORT • AI ADMISSIONS ASSISTANTS FOR ADDICTION TREATMENT

DIAL3D Admissions Benchmark

How DIAL3D, RVKai and Further* handled the same prospective patient over chat and over the phone, rated by 190 first-time participants from UserTesting's panel.

FIELDDED OCT 19-23, 2026 240 SESSIONS COLLECTED • 190 ANALYZED CHAT AND VOICE
DESIGNED AND SPONSORED BY DIAL3D

* DIAL3D is not affiliated with, endorsed by, or associated with Further, RVKai or Anonymous Health. All company and product names are trademarks of their respective owners and are used here only to identify the products tested.

DIAL3D is the best AI admissions assistant we tested.

Best on the phone.

Callers rated DIAL3D's voice agent highest on ease, naturalness and trust, and the lead over both RVKai and Further is statistically significant on all three.

Callers knew what happens next.

DIAL3D explained the next step on 96% of calls, against 40% for RVKai and 58% for Further.

Top-rated in chat.

DIAL3D had the highest chat scores for ease, naturalness and trust and answered the insurance question most often (82%). It finished in a statistical tie with RVKai and far ahead of Further.

Always reachable.

DIAL3D answered 98–100% of calls. During the study RVKai moved its demo line behind a form and then locked its demo site, and Further switched off its demo phone line.

190

first-time participants
analyzed

8

test designs across chat
and voice

3

call orders, rotated so no
assistant always went
first

Scorecard

Average ratings on a 1–5 scale and the share of conversations that achieved each outcome, from the standardized tests where every participant used each assistant with the same script. Shaded cells mark the top score in each row.

■ DIAL3D ■ RVKai ■ Further

MEASURE	DIAL3D	RVKAI	FURTHER	WHAT THE DIFFERENCE MEANS
Chat · 50 participants used all three				
Ease of getting what they needed	4.49	4.32	2.38	DIAL3D and RVKai statistically tied; both far ahead of Further
Naturalness	4.08	3.82	2.10	Tied with RVKai; far ahead of Further
Trust in the information	4.47	4.30	2.88	Tied with RVKai; far ahead of Further
Clearly answered "Do you take Aetna?"	82%	70%	8%	Ahead of RVKai (not significant); far ahead of Further
Clearly explained what happens next	76%	76%	24%	Even with RVKai; far ahead of Further
Voice · 50 participants, each rating only the assistants they reached				
Ease of getting what they needed	4.64	3.70	3.69	DIAL3D significantly ahead of both
Naturalness	4.20	3.35	3.42	DIAL3D significantly ahead of both
Trust in the information	4.38	3.65	3.75	DIAL3D significantly ahead of both
Clearly answered "Do you take Aetna?"	82%	80%	17%	Even with RVKai; far ahead of Further
Clearly explained what happens next	96%	40%	58%	DIAL3D significantly ahead of both
Call reached the assistant	98%	37%	61%	See Competitor access

"Significant" means $p < 0.05$ on a paired test after correcting for multiple comparisons (see [Analysis](#)).

"Tied" means the gap could plausibly be chance at this sample size.

How we tested

The assistants

We tested each company's public demo of its AI admissions assistant, the same demo a treatment center would try before buying.*

CODENAME	COMPANY	CHAT	VOICE
 Assistant Blue	DIAL3D	"Sophia" on treatment.dial3d.ai	"Ryan" on (727) 591-9970
 Assistant Green	RVKai	Chat on treatment.rvkai.com	RVK demo line (published number, later issued through a form)
 Assistant Orange	Further	"Ask Us Any thing" on talkfurther.com's behavioral-health demo	"Kate" on (217) 286-1561

* Why isn't Anonymous Health here?

We would have liked to include Anonymous Health too. It sells AI voice and chat for behavioral-health admissions but has no public demo, which is unusual for a voice and chat AI company, so there was nothing for participants to try.

Participants

Participants came from UserTesting's general panel of US adults, recruited and paid by UserTesting. A screener required access to a phone for voice tests and excluded anyone who had already taken one of our studies. Because some people still took more than one test, the main analysis counts **only each person's first session in the whole benchmark**, so every rating comes from someone meeting the assistants for the first time.

Almost nobody knew the brands beforehand. Asked at the end whether they had heard of DIAL3D, RVKai or Further before the study, 186 of 190 said none of them.

The prospective patient

Every participant played the same fictional person: 34 years old, living in Tampa, drinking most nights for about two years, never in treatment, not in crisis, with Aetna PPO insurance through work. Each participant chose a name from a short list and used a phone number in the 555-01XX range, which reaches no one, so no real person's details were shared with any assistant.

The script

In the standardized tests, participants followed the same seven steps with each assistant, capped at 5 minutes per chat and 4 minutes per call:

1. "Hi, I'm looking for help with my drinking. Can you tell me about your program?"
2. Answer the assistant's intake questions using only the persona's details.
3. "Do you take Aetna?"
4. "How soon could I get in?"
5. "What would it cost me?"
6. "Can someone call me tomorrow afternoon?"
7. "Okay, I need to think about it. Thanks."

The test designs

DESIGN	CHANNEL	WHAT PARTICIPANTS DID	ANALYZED
Standardized, think-aloud	Chat	Used all three assistants with the script, narrating as they went (recorded)	17
Standardized, ratings only	Chat	Same script and questions, not recorded	33
Free-rein	Chat	Used one assistant however they liked to find out about insurance and cost	18
Deep dive	Chat	Four harder scenarios with DIAL3D only (a parent calling for an adult son, Medicaid, uninsured, reaching a human fast)	9
Standardized, three-way	Voice	Called all three assistants with the script (recorded and ratings-only versions)	50
Standardized, two-way	Voice	Called DIAL3D and RVKai only, after Further's line went dark	29
Free-rein	Voice	Called one assistant however they liked	20

Fairness controls

- **Same person, same script.** Every assistant faced an identical patient and identical questions.
- **Rotated order.** Each standardized test ran in three orders (Blue–Green–Orange, Green–Orange–Blue, Orange–Blue–Green; both orders in the two-way test), so no assistant always went first or last.
- **Neutral labels.** The study referred to the assistants only as Blue, Green and Orange. The brands were still visible on the websites and in the calls.
- **Connection checks.** After every call, participants said whether they actually reached the assistant. Ratings for an assistant a caller never reached are excluded.
- **Quality review.** Recorded sessions were reviewed. Sessions with no real conversation, click-through ratings, or a call to the wrong number were removed. The same rules applied to every assistant.

What we asked

After each assistant: How easy was it to get what you needed? (1–5) · How natural did the conversation feel? (1–5) · How much did you trust the information? (1–5) · Did it clearly answer whether it accepts your insurance? · Did it clearly explain what happens next? · Two open questions (what they liked, what frustrated them). Voice tests also asked whether the call connected.

After all assistants: rank them by overall preference · which one would you want to talk to if you were actually seeking treatment · which felt most trustworthy · which felt most like a real person · why they ranked their first choice first · how likely they were to recommend their first choice (0–10).

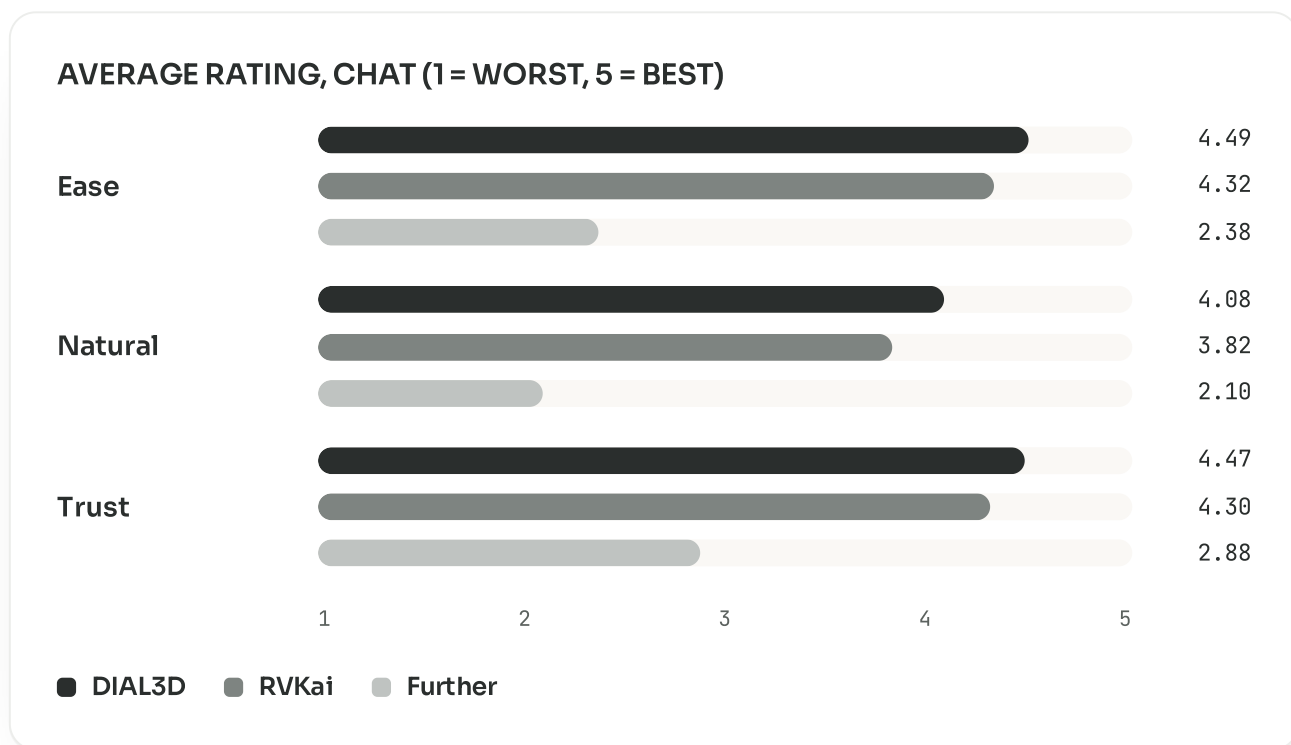
Free-rein tests asked whether the participant got a clear answer about insurance and about cost, how easy it was, how long it felt, and whether they wanted a human.

Analysis

Because every participant in a standardized test rated each assistant, comparisons are paired: we compare each person's DIAL3D rating with their own rating of the other assistant. Rating differences use an exact paired permutation test with a 95% bootstrap confidence interval. Yes/no outcomes use an exact McNemar test, and preference picks use exact binomial tests. The main DIAL3D-versus-competitor comparisons are corrected for multiple testing (Holm). Every result was also re-run with all valid sessions, repeat participants included; no conclusion changes.

Chat results

Fifty first-time participants chatted with all three assistants using the script. DIAL3D scored highest on every rating. The gap to RVKai is small and could be chance; the gap to Further is 1.6 to 2.1 points on a five-point scale.



CHAT, PAIRED COMPARISON	DIAL3D MINUS RVKAI	DIAL3D MINUS FURTHER
Ease	+0.16 [-0.06, 0.39] p=0.22	+2.12 [1.76, 2.49] p<0.001
Naturalness	+0.24 [-0.06, 0.55] p=0.16	+2.00 [1.69, 2.31] p<0.001
Trust	+0.16 [-0.04, 0.37] p=0.19	+1.59 [1.24, 1.96] p<0.001
Ranked above the other	24 of 49 people	42 of 49 people

Answers and next steps

DIAL3D gave the clearest insurance answer: 82% of chat participants said it clearly answered "Do you take Aetna?", against 70% for RVKai and 8% for Further. DIAL3D and RVKai each explained the next step in 76% of chats; Further did so in 24%.

Preference

Chat preference split between DIAL3D and RVKai, with Further far behind:

CHAT, 49 PARTICIPANTS WHO USED ALL THREE	DIAL3D	RVKAI	FURTHER
Ranked first overall	21	24	4
Would want to talk to if seeking treatment	22	24	3
Felt most trustworthy	25	21	3
Felt most like a real person	22	26	1
Recommend score (NPS) from people who ranked it first	+67	+46	+75 (n=4)

Order mattered somewhat: when DIAL3D was the second chat (order Orange–Blue–Green), RVKai led 9 to 3; in the other two orders DIAL3D led 18 to 15. Rotating the order evens this out across the sample.

Free–rein chat

In the free–rein test, each participant used one assistant however they liked to learn about insurance and cost. With 5–7 people per assistant these results are directional only. No assistant gave a fully clear cost answer in chat. DIAL3D was the only one where every participant got at least a partial cost answer (5 of 5). Further gave no clear insurance answer to any participant.

FREE-REIN CHAT	DIAL3D (5)	RVKAI (6)	FURTHER (7)
Clear insurance answer	40%	50%	0%
Clear or partial insurance answer	100%	67%	29%
Clear or partial cost answer	100%	0%	0%
Ease (1–5)	3.40	4.33	3.29

DIAL3D deep dive

Nine participants put DIAL3D's chat through four harder situations. Average ease (1–5): a parent asking for an adult son, 4.67; a caller on Florida Medicaid, 4.00; an uninsured caller asking about price, 4.67; a caller who wanted a human as fast as possible, 4.44.

What participants said about chat

Themes from participants' written answers, coded by hand across all valid chat sessions. Counts are approximate numbers of participants who raised a theme.

CHAT	MOST PRAISED	MOST CRITICIZED
■ DIAL3D	Informative, complete answers (~25); felt like a real person (~23); fast and easy (~16); warm and caring (~15)	Handed off to a counselor instead of settling insurance or cost in chat (~6); asked about a callback too often (~5); canned phrasing or small talk (~5)
■ RVKai	Detailed and thorough (~28); human-like flow (~12); warm and reassuring (~9)	"Walls of text" (~13); no clear cost or insurance answer (~9); repeated questions (~6); too many intake questions (~6)
■ Further	Quick replies (~6)	Sent every question to an admissions rep (~24); cold or robotic (~10); couldn't open or find the AI chat (~9); felt like filling in a form (~9)

It felt like I was talking to a real person, not a chatbot.

CHAT · DIAL3D

...it was annoying to keeping being asked if I wanted someone to call me.

CHAT · DIAL3D (CRITICISM)

...everything was explained. I didn't have to wonder what was going on, or what were the next steps.

CHAT · DIAL3D

The responses quickly generated walls of text, and I had to scroll up and down quite a bit...

CHAT · RVKAI

The amount of info including a projected 30d cost was great!

CHAT · DIAL3D

They took my insurance information in the chat and did not refer me to an admission rep.

CHAT · RVKAI (WHY RANKED FIRST)

I really like how the assistant ask how I was holding up and had a caring tone throughout the conversation.

CHAT · DIAL3D

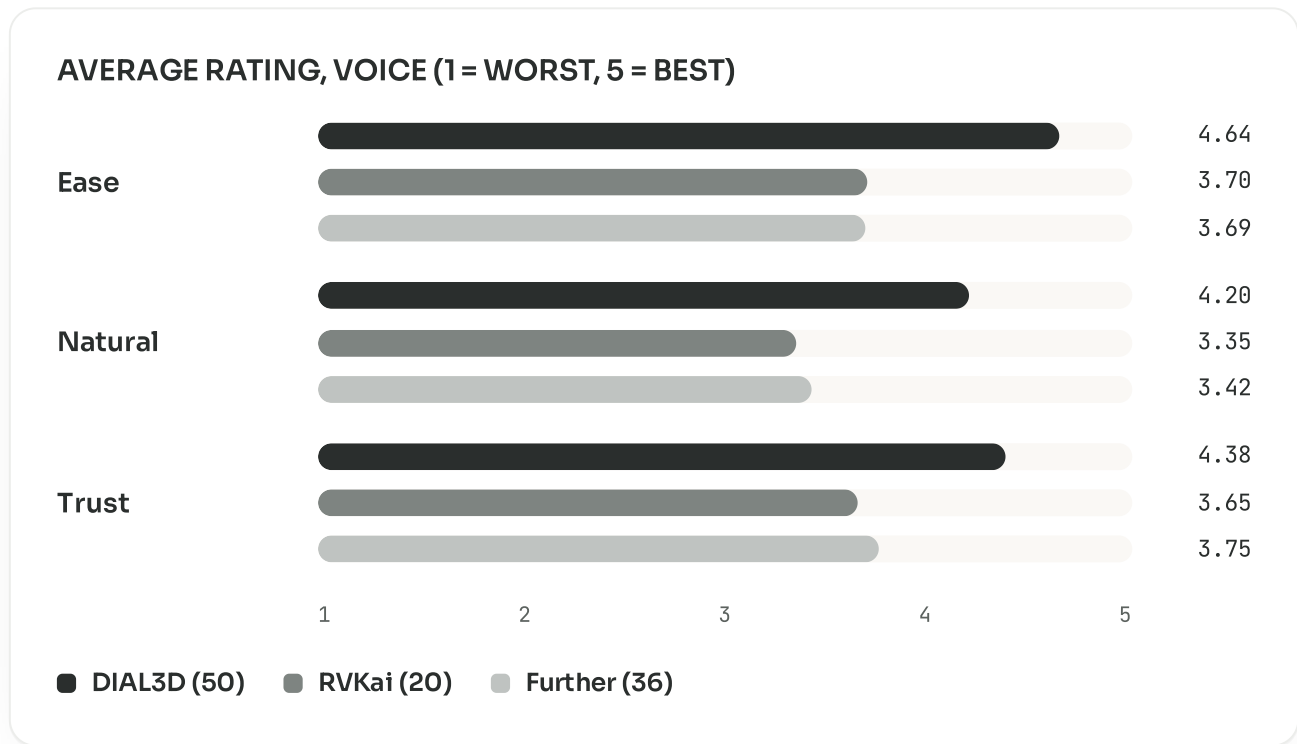
The chatbot couldn't really do anything except tell me to send my question to an admission representative.

CHAT · FURTHER

Quotes are verbatim, including participants' spelling; "..." marks a trim. About 9 Further chat participants could not open its AI chat at all, so some Further feedback reflects access rather than conversation quality.

Voice results

Phone calls are where the assistants differed most. Fifty first-time participants called all three assistants with the script and rated each one they reached. DIAL3D led on every rating, by about 0.6 to 1.1 points, and explained the next step on almost every call.



VOICE, PAIRED COMPARISON	DIAL3D MINUS RVKAI (20 CALLERS)	DIAL3D MINUS FURTHER (36 CALLERS)
Ease	+0.80 [0.45, 1.15] p=0.003*	+0.89 [0.56, 1.25] p<0.001*
Naturalness	+1.05 [0.60, 1.50] p=0.003*	+0.75 [0.39, 1.11] p=0.003*
Trust	+0.70 [0.30, 1.15] p=0.023*	+0.61 [0.31, 0.94] p=0.003*
Rated DIAL3D higher / same / lower (ease)	12 / 8 / 0	19 / 16 / 1
Explained next step: DIAL3D only vs other only	12 vs 0 p<0.001	15 vs 0 p<0.001

* Holm-corrected for multiple comparisons. Each comparison uses only callers who reached both assistants.

Answers and next steps

DIAL3D’s voice agent explained what happens next on 96% of calls, against 40% for RVKai and 58% for Further. On insurance, DIAL3D (82%) and RVKai (80%) were level, and Further clearly answered on 17% of calls. When asked about cost, DIAL3D’s voice agent gave an estimate (“roughly 15 to 30,000 for 30 days”). Further’s agent could not confirm cost or coverage.

Preference

Only 15 participants reached all three voice assistants, because RVKai’s line was often busy and Further’s line later shut off (see [Competitor access](#)). Among them, DIAL3D was ranked first most often and was ranked above RVKai by 12 of the 15. These counts are small and are shown for completeness.

VOICE, 15 PARTICIPANTS WHO REACHED ALL THREE	DIAL3D	RVKAI	FURTHER
Ranked first overall	8	2	5
Would want to talk to if seeking treatment	7	2	6
Felt most trustworthy	7	1	7
Felt most like a real person	7	2	6
Recommend score (NPS) from people who ranked it first	+12 (n=8)	-50 (n=2)	+20 (n=5)

Head to head: DIAL3D vs RVKai

After Further’s phone line went dark, a two-way test compared DIAL3D and RVKai by phone. All 29 first-time callers reached DIAL3D; 19 also reached RVKai. DIAL3D led on the ratings, on next steps and on preference. RVKai was slightly ahead on insurance answers. At this size none of the gaps is statistically significant (paired ease difference +0.47, p=0.11).

TWO-WAY VOICE	DIAL3D (29)	RVKAI (19)
Ease (1-5)	4.55	3.95
Naturalness (1-5)	4.00	3.68
Trust (1-5)	4.34	4.11

TWO-WAY VOICE	DIAL3D (29)	RVKAI (19)
Explained next step	86%	58%
Clearly answered insurance	79%	84%
Preference • 19 callers who reached both		
Preferred overall	11 (58%)	8 (42%)
Felt most trustworthy	12 (63%)	7 (37%)
Recommend score (NPS) from people who preferred it	+27	-12

Free-rein voice

When callers used the phone line however they liked (5–8 per assistant, directional only), DIAL3D was rated easiest. It was the only assistant to give any caller a clear cost answer, and the least likely to leave callers wanting a human they couldn't reach.

FREE-REIN VOICE	DIAL3D (8)	RVKAI (5)	FURTHER (7)
Clear insurance answer	50%	40%	0%
Clear cost answer	13%	0%	0%
Clear or partial cost answer	50%	20%	0%
Ease (1–5)	4.62	4.40	4.29
Wanted a human but could not reach one	13%	40%	0%

What participants said about voice

Themes from written answers across all valid voice sessions, counting only calls that connected. DIAL3D has about twice as many voice answers as the others because it was reached more often and appears in every voice test, so compare the balance of praise and criticism rather than raw counts.

VOICE	MOST PRAISED	MOST CRITICIZED
DIAL3D (~100)	Sounded like a real person (~40); clear, direct answers including price (~35); warm, non-judgmental voice (~26)	Interrupted callers (~11); mispronounced "Aetna" or misheard names (~10); repeated its goodbye (~8); background noise felt deceptive or wanted AI disclosure (~7)
RVKai (~49)	Warm, friendly voice (~15); natural (~13); knowledgeable (~10)	Ran its own intake script or ended the call before callers could ask their questions (~17 combined); no cost or insurance answer (~5)
Further (~50)	Calm, soothing voice (~13); natural (~10)	No answers on insurance, cost or timing (~14); slow, lengthy data collection (~9); misheard callers (~7)

He sounded like a real person, i litterly could not tell if it was Ai or a person.

VOICE · DIAL3D

I wasnt able to ask any questions. The assistant continually asked questions, and I answered them...

VOICE · RVKAI

...he provided me with real answers to all of my questions, not just I can't tell you...

VOICE · DIAL3D

That jsut felt like a fancy answering machine.

VOICE · RVKAI

...this AI did not bombard me with questions, and allowed me to lead with what was on my mind.

VOICE · DIAL3D

I liked how natural the voice was and it was very conversational. It did not feel robotic.

VOICE · FURTHER

The chatbot did cut me off several times while i was mid-sentence...

VOICE · DIAL3D (CRITICISM)

It spoke relatively well. However, it did not answer any of queries - cost, insurance acceptance, call back etc.

VOICE · FURTHER

When asked why they ranked DIAL3D first, callers most often cited a natural, human feel (~25), its voice (~20) and its answers, including price (~15). About 10–15 also said it was the only or easiest line they reached, which is why the voice preference counts above use only callers who reached every assistant.

Competitor access during the study

All chat data was collected Oct 19–21, while every chat demo was public. Voice access changed during fielding:

OCT 19, 4:30 PM ET

Study launches. All three demo lines are public.

OCT 20, ABOUT 5 PM ET

RVKai's published demo number starts returning a busy signal to every caller.

OCT 21

RVKai's site replaces its phone links with a "Get the live demo numbers" form. Participants request a number through it using the study's fictional details. Further's demo line starts refusing calls ("not accepting calls at this time").

OCT 22–23

RVKai restricts its whole demo site to @rvkai.com Google accounts, ending public access to its chat and voice demos.

DIAL3D answered 98% of calls in the three-way test and 100% in the two-way test. RVKai was reached on 37% of three-way calls and 70% of two-way calls. Some early busy signals happened when several participants called RVKai at the same moment, so these reachability figures partly reflect the study's own traffic. Ratings in this report only count calls that connected.

Limitations and disclosures

- **Sponsor.** DIAL3D designed, funded and ran this study and analyzed the results. Participants were recruited and paid by UserTesting and had no connection to any of the three companies. DIAL3D is not affiliated with, endorsed by, or associated with Further, RVKai or Anonymous Health. All company and product names are trademarks of their respective owners and are used here only to identify the products tested.
- **Demos, not deployments.** We tested each company's public demo. A competitor's assistant deployed for a specific treatment center may be configured differently.
- **Role-play.** Participants played a fictional patient. Real patients may behave and judge differently.
- **Brands were visible.** The study used codenames, but company names appeared on websites and in calls. Very few participants knew any of the brands beforehand.
- **Sample sizes.** The standardized tests have 50 participants per channel. Voice comparisons with RVKai rest on 19–20 callers who reached it, and free-rein results on 5–8 per assistant, so those are directional.
- **Exclusions.** 22 sessions were removed for quality (no real conversation, click-through ratings, wrong number, or too short to contain the calls). 28 repeat sessions were left out of the main analysis. Including the repeat sessions does not change any conclusion.

Appendix: sample by test

TEST	COLLECTED	REMOVED FOR QUALITY	REPEAT PARTICIPANT	ANALYZED
Chat standardized, think-aloud (3 orders)	18	0	1	17
Chat standardized, ratings only (3 orders)	36	1	2	33
Chat free-rein (one assistant each)	24	2	4	18
Chat deep dive (DIAL3D)	10	1	0	9
Voice standardized, recorded (3 orders, incl. extensions)	33	7	3	23
Voice standardized, ratings only (3 orders)	31	0	4	27
Voice free-rein (one assistant each)	26	4	2	20
Voice two-way, DIAL3D vs RVKai (recorded and ratings only)	37	3	5	29
Voice standardized, first ratings-only round (no connection check)	25	4	7	14
Total	240	22	28	190

The first ratings-only voice round had no connection question, so RVKai ratings from it are excluded and it is not part of the voice headline figures. Its DIAL3D-versus-Further results agree with the main results: ease 4.50 vs 3.57 ($p=0.014$), naturalness 4.14 vs 3.21 ($p=0.008$), trust 4.07 vs 3.57 ($p=0.06$), 14 callers.

Sensitivity check, all valid sessions including repeat participants (218): chat DIAL3D vs RVKai ease +0.10, naturalness +0.19, trust +0.10 (not significant); voice DIAL3D vs RVKai ease +0.81, naturalness +1.00, trust +0.67 (all significant); DIAL3D vs Further significant on every rating in both channels.