CHEAT SHEET · SYNTHETIC RESEARCH

Synthetic personas & digital twins: *what they are good for*

FlashInsight
Dr Laurent Florès · flashinsight.io
laurent@flashinsight.io

THE ONE-MINUTE VERSION

A caricature, not a portrait

Recent published tests: right direction, gaps between groups too wide, disagreement inside a group erased. Good for recognising and comparing, not for measuring.

Upstream, not instead

Explore and filter before fieldwork. Keep real people for thresholds, breakthroughs and emotion.

Against people, not against 100%

Real people asked twice do not answer the same way every time. Good research practice: read any accuracy claim against that, never against 100%.

WHICH TOOL FOR WHICH QUESTION

You want to...	Use
Challenge what you think you know about your consumers	**A PersonaBot: your knowledge, given a voice**
Explore many options, targets or markets before fieldwork	**Synthetic personas**
Go deeper on one well-documented profile	**A digital twin (today, it answers like its type)**
Set a threshold, judge a breakthrough, read emotion, decide go/no-go	**Real people**

✓ Works well

- Exploring a direction before fieldwork
- Pre-testing a concept, an ad or a wording; exploring many options
- Contrasting segments and markets, including those the budget did not cover
- Building hypotheses to take to real people

✕ Not yet

- Standing in for one particular person
- Fine differences inside a segment, fine cross-tabs, a ranking of close options
- Absolute levels: percentages, market sizing, "launch if 40% buy"
- Genuinely new ideas and emerging trends, cultural nuance, emotion
- What people know or have noticed (awareness)

BEFORE YOU RUN: FOUR CHECKS

1. **Write the decision down.** A comparison between options, or a threshold? A threshold goes to people.
2. **Keep a human reference at hand.** A past study on a similar question shows where synthetic answers agree and where they drift.
3. **Describe your target in layers:** who they are, what they do and, when you have it, what they value. Values often help explain the why: check it against real data.
4. **Plan the human step** from the start: what you will check with real people, and when.

READING THE RESULTS: FIVE RULES

1. Read directions and contrasts, not percentages, and confirm them with people.
2. Trust large, consistent gaps; treat small gaps and close rankings as noise.
3. Be wary of a consensus that is too perfect: real people disagree.
4. The newer the idea, the more weight goes to people.
5. Compare against people asked again, never against 100%.

FIVE QUESTIONS TO ASK ANY SUPPLIER, US INCLUDED

Ask	A good answer
Data. What are the personas built on?	Named sources (public statistics, your data), and whether attitudes and values are included
Validation. Validated against which real people?	A named human study, a human benchmark (two samples, or the same people asked again), and revalidation when the model changes
Limitations. Does it keep the disagreement inside a segment? What if information is missing?	Spread measured and reported; a model that can say "I don't know"
Fit for purpose. Is it fit for my decision?	Directions and contrasts yes, thresholds no; they tell you when to go to people
Transparency. Will the report say what was simulated?	Yes: synthetic outputs labelled apart from human data, human oversight stated (ICC/ESOMAR Code, 2025)

A FIRST PILOT IN FOUR STEPS

1. Pick one past decision with known human results: a concept, claim or ad test, one market.
2. Run the same stimulus, questions and target with synthetic personas, blind to the human results.
3. Compare: same direction? Same large contrasts? Where does it drift?
4. Decide where it helps you: upstream, to explore and filter. People for the rest.

Write down what "good enough" means before you look: for example the same direction on the large contrasts.

Want to try it on one of your own decisions?
flashinsight.io/en/contact

More, and the studies behind this sheet: flashinsight.io

Sources: Pew Research Center (2026); Peng et al., Science Advances (2026); Brand, Israeli & Ngwe (2025); Park et al. (2024); ICC/ESOMAR Code (2025); ESOMAR, 10 Questions for buyers of AI-based research (2026).