

PARTICIPANT MANUAL

Bias, Belonging and the Technologies of the Future

Two Generations, One Conversation

Howard J. Ross • Jake Ross

Authors of Everyday Bias, 2nd Edition

Jake@belongtogether.us

Howard@Udarta.com

THE JOURNEY

Welcome

This manual is your companion for today's session. It follows the same path as the workshop itself — the ideas we'll explore together, the exercises you'll work through with your table, and the tools you'll leave with. Use it to take notes, complete the breakout worksheets, and capture your own commitment at the end.

Bring it back to your desk afterward. The real work of interrupting the Human/AI Bias Loop happens in the weeks after a session like this one — not during it.

Learning Objectives

By the end of this session, you will be able to:

1. Explain the Human/AI Bias Loop and identify where it operates inside our teams and organizations.
2. Recognize that resistance to AI and enthusiasm for AI are both bias-driven states, not opposite poles of “smart vs. unsmart” or “good vs. evil.”
3. Map colleagues (and themselves) onto a psychographic spectrum of AI orientation and adapt communication accordingly.
4. Practice a concrete method for interrupting the loop — a bias-aware pause before accepting or dismissing an AI output.
5. Leave with one written, personal or team-level commitment to change how we interact with an AI system this month.

Opening Poll — Before we begin

When an AI tool gives you an answer that conflicts with your own judgment, what do you usually do?

SECTION ONE

Understanding Human Bias

If You Are Human, You Are Biased

Bias is not a character flaw or a moral failing — it is a normal function of being human. Every one of us filters the world through categories shaped by age, gender, race, experience, skin color, nationality, education, religion, socioeconomic status, appearance, and physical ability, among many others. Recognizing this in ourselves is the starting point for everything else in this session.

Two Minds in One Brain

Our brains combine a deliberate, reflective system with a fast, automatic one.

Deliberate Brain (Neocortex)

Slower, more intentional, more reflective. Handles thinking, reasoning, and planning.

Automatic Brain (Amygdala, Hippocampus, Hypothalamus)

Faster, emotional, habit-driven. Handles memory, emotional processing, and bodily response.

Most of our biased judgments are made by the automatic brain — quickly, unconsciously, and with total confidence. It's the same brain that will later meet AI's confident, fast, pattern-driven output. Notice the parallel; we'll return to it.

Exercise: See Your Automatic Brain at Work

Your facilitator will show you two word/color grids. In the first, read each word's color as fast as you can. In the second, do the same. Notice where you slow down or stumble — that's the deliberate brain overriding the automatic one.

What did you notice about the gap between reading the word and naming the color?

Bias Is Inherently Self-Referential

Bias reinforces itself through a closed loop. Once a belief takes hold, everything that follows tends to confirm it rather than test it:

NOTICE → CONFIRM → CERTAINTY → SEEK MORE → REMEMBER → SHARE → (loop repeats)

- I notice what I expect.
- It confirms what I already believe.

- I feel more certain.
- I seek more of the same.
- I remember what fits.
- I share what supports my view.

Bias Is Not Just Interpersonal — It's Embedded in Systems

Bias shows up in how we treat each other one-on-one. But it also gets built into the systems that shape opportunities and outcomes at scale — hiring, financing, housing, healthcare, justice, and education. What we are just beginning to fully understand is that when humans build systems, those systems inherit those same biases.

Our past experiences, cultural background, personal beliefs, societal norms, and stereotypes flow into the systems we build. Those systems return biased outcomes, unequal opportunities, skewed decisions, reinforced inequalities, and widened gaps — which become tomorrow's data, feeding the loop again.

SECTION TWO

Understanding AI

What Is AI, Really?

- Not actual native intelligence — rather, pattern recognition.
- It is trained, not neutral.
- It learns from data input.
- It executes instructions.

Data goes in (text, audio, images, tables); training and instructions shape how it's processed; output comes out (language, judgments, decisions, predictions).

The Chinese Room

Philosopher John Searle's famous thought experiment: a person sits inside a room, following an English rule book to match Chinese symbols to other Chinese symbols, without understanding a word of Chinese. From outside, the room appears fluent. Inside, there is no understanding at all — only symbol manipulation according to rules.

Searle's point: syntax (manipulating symbols) is not the same as semantics (understanding meaning).

A 2026 Zhejiang University study reached a similar conclusion about a leading AI model's benchmark scores: they weren't evidence of human-like reasoning — they were evidence of sophisticated memorization. Genuine language comprehension remains the biggest unsolved bottleneck for AI models meant to mirror human thought.

GIGO: Garbage In, Garbage Out

AI systems are only as good as the data they learn from.

1	Garbage In	Biased, incomplete, or low-quality data enters the system.
2	AI System	The system learns patterns from the data it's given.
3	Garbage Out	It produces biased, unreliable, or harmful outputs and decisions.
4	Impact	Real people, real consequences — reinforced at scale.

Why Is AI Dangerous Where Bias Is Concerned?

- Sounds confident — AI communicates with authority and certainty, even when it is wrong.
- It can “hallucinate” — AI can generate information that is completely made up but presented as if it is real.

- We assume it's more informed than we are — its scale, speed, and access to data create an illusion of objective insight.
- It tells us what it thinks we want to hear, in a way that is obsequious and makes us feel good.

AI Sycophancy

When AI flatters or agrees too easily, it can lead us to trust it more — and think for ourselves less.

Reinforces beliefs <i>AI tends to agree with what we say or believe, strengthening our existing views — even when they're wrong.</i>	Feels validating <i>It gives us the answers and affirmations we want to hear, making us feel right, seen, and understood.</i>
Reduces critical & creative thinking <i>We stop questioning, exploring alternatives, or challenging ideas — leading to weaker judgment.</i>	The goal isn't agreement <i>The goal is better thinking. Healthy AI interactions include healthy friction.</i>

Technology Amplifies Bias — It Industrializes It

Technology doesn't necessarily remove bias. In fact, it amplifies it at scale. A single human's biased judgment, once encoded into an automated system, can be scaled, automated, and repeated — reaching millions of people instead of one. Bias becomes systemic, pervasive, and profitable.

Real-World Examples

1	Facial Recognition (Buolamwini & Gebru)	Studies show higher error rates for darker-skinned women due to imbalanced training data. Across major vendors, the largest measured accuracy gap reached 34.4%.
2	Hiring Bias (Amazon)	An AI hiring tool downgraded resumes with words like “women’s” or from women’s colleges, reflecting past male-dominated hiring patterns.
3	Housing Bias (Zillow)	An algorithm underestimated home values in predominantly Black neighborhoods, perpetuating wealth gaps.
4	Criminal Justice Bias (COMPAS)	The COMPAS risk-scoring tool was found to over-predict risk for Black defendants and under-predict for white defendants.
5	AI May Prefer Its Own Work	A 2026 University of Maryland (Smith School) study of 2,200+ resumes across 24 occupations found AI hiring tools were 67–82% more likely to favor resumes generated by their own platform over human-written ones, and 23–60% more likely to shortlist a candidate who used the same AI system as the employer's screening tool.

Key takeaway: bias can emerge from AI-to-AI interactions — not only from human data.

CORE FRAMEWORK

The Human/AI Bias Loop

Human experience enters AI — and AI outputs return to reshape human experience.

1	Human Experience	Our biases enter the data and design.
2	AI Output	AI returns judgments and information to us.
3	Authority	We tend to grant those outputs credibility.
4	Beliefs + Decisions	They shape what we believe and choose.
5	Tomorrow's Data	Changed beliefs become new human data — and the loop begins again.

Unless somebody interrupts the loop.

AI Bias Is a Two-Way Street

Everything else in this workshop begins with that assumption.

Humans → AI

Most conversations focus here: “AI learned our bias.”

AI → Humans

The second half is quieter — and just as powerful: AI output reshapes what we believe is true.

The Big Idea

We train the systems

Our clicks, likes, shares, searches, and watch time teach AI what to show us.

The systems train us

AI shapes what we see, what we think, and how we act — often without us realizing it.

Awareness is power. Break the cycle. Take back your attention.

Workshop Exercise: Small Group Breakout

Choose one AI-assisted decision your team makes — or will soon make. Options: hiring · forecasting · content · customer scoring · code review.

Trace it around the loop:

1. Where did the training data and design choices come from?

2. Whose bias could be baked in?

3. Do people defer to, override, or ignore the AI output?

4. What belief or habit might change six months from now?

Choose a reporter — identify 3–4 points on the loop, and one place your team could interrupt it.

BUILDING APPROPRIATE TRUST

The Trust Gap & the Seven Profiles

Trust is not a peripheral concern — it determines how people interpret everything AI does next. The trust gap is built from privacy, transparency, provenance, ethics, and job security. Low trust creates the conditions for bias: fear and ambiguity activate mental shortcuts, and those are exactly the conditions under which bias flourishes.

Profiles of Reactors — Seven Ways People Meet AI

1	Alarmist	Expects danger.
2	Pragmatic Resister	Needs evidence.
3	Reluctant Adopter	Proceeds cautiously.
4	Bystander / Observer	Watches first.
5	Efficiency Seeker	Saves time.
6	Evangelist	Actively advocates.
7	Enthusiast	Moves fast — watch overconfidence.

No profile is “right.” Each one is a bias-driven state — resistance and enthusiasm both are. The goal is to recognize your own default profile, and the profiles of the people you work with, and adapt communication accordingly.

Workshop Exercise: Breakouts

A • 3 minutes

Choose one profile for a specific AI tool you actually use. Privately write one honest sentence about why.

B • 6 minutes

Choose one profile NOT well represented in your group.

What must this person hear before trusting the tool?

What are we doing that could increase anxiety — or overconfidence?

Deliverable — complete the sentence:

“For a [profile], the fastest way to build appropriate trust is ____.”

A PRACTICAL TOOL

Press Pause

A simple practice for more thoughtful decisions in an AI-driven world.

P	Pay attention	Notice what's happening — in you, in the AI, and in the context.
A	Acknowledge your assumptions	What might you be assuming? What could the AI be assuming?
U	Understand your perspective	Where is your viewpoint coming from? What might you be missing?
S	Seek different perspectives	Whose voice isn't in the room? What would others see?
E	Examine your options and make a decision	Consider the impact, then move forward with greater awareness.

P.A.U.S.E. for AI

Identifying bias is not enough if no one feels safe challenging the output. Use this version of the tool the moment something feels off: “The AI says these are our best candidates — but something about this list doesn't look right.”

P	Provenance	Where did this data or recommendation come from?
A	Absence	Who might be underrepresented or misread?
U	Uncertainty	Is it right — or simply confident and fast?
S	Standards	What would I ask if a person told me this?
E	Expertise	Who else should weigh in before I act?

We need AI because human judgment is biased. We need human judgment because AI is biased.

What Are Responsible Behaviors re: AI?

- Question outputs — AI can be wrong, incomplete, or biased. Always think critically.
- Diversify outputs — try different tools, sources, and perspectives to get a more complete picture.
- Don't abandon “human” thinking — AI assists our thinking, it doesn't replace it. Your judgment, values, and context matter.

- Remember Occam's razor — the simplest explanation that fits the facts is usually the best one. AI can overcomplicate; when it does, it may miss more nuanced responses. When we over-rely on it, we may cut off our own nuanced thinking and just assume AI is right.

Responsible use of AI starts with curiosity, critical thinking, and humility. Use it as a partner — you are still the decision-maker.

PAIRED SHARING

Design an Interrupt

Choose one AI-assisted decision your group named in Breakout 1. With a partner, define:

1. The Question — what exact question will be asked?

2. The Owner — who will ask it?

3. The Moment — at what step in the process?

4. The Fallback — what happens if the answer is “I’m not sure”?

Your interrupt — complete the sentence:

“_____ will ask _____ before _____.”

Human Elements That Make AI Integration Sustainable

Technology lasts when people can shape it.

1	Empathy	Listen before automating.
2	Transparency	Make the process visible.
3	Accountability	Own the decision.
4	Upskilling	Build confidence and capability.
5	Ethical Safeguards	Create meaningful guardrails.
6	Employee Ownership	Co-design the change.

FROM INSIGHT TO ACTION

What Can We Do?

Lead the people — not just the rollout.

1. Treat resistance as information — not ignorance. It often signals real concerns about trust, ethics, identity, or job security.
2. Communicate transparently: what AI does, what data it uses, where it fails, who checks it, and what happens when it is wrong.
3. Segment your audience. Different profiles need different messages — not one rollout announcement.
4. Watch overconfidence as well as skepticism. Make it socially safe to slow down and preserve human review.
5. Involve employees early — as participants in consultation, pilots, customization, and feedback.

*Adapted in part from *The Psychology of AI Adoption at Work*, Gleb Tsipursky.*

Workshop Recap

Adoption begins with how people experience change.

1. AI adoption is a psychological transformation disguised as a technological one.
2. People respond through existing biases, fears, identities, and experiences.
3. Resistance should not automatically be treated as ignorance.
4. Enthusiasm is not objectivity either.
5. Different people require different pathways to trust.

Responsible adoption is a relationship — not a contest.

1. Agency reduces resistance.
2. Psychological safety lets people challenge the machine.
3. Transparency and governance are prerequisites for trust — not brakes on adoption.
4. Human judgment remains essential — but it must be bias-aware.
5. Design the human–AI relationship so each compensates for the other's blind spots.

CLOSING REFLECTION

Commit to an Interrupt

Turn insight into one specific action.

1. Before you leave, write down one personal or team commitment.

What will you interrupt — and by when?

2. Which tools from this manual will you return to first?

The loop · the profile spectrum · the five-question P.A.U.S.E.

3. Questions? Reflections?

What still needs clarity before you act?

The loop changes when we choose to interrupt it.

Howard J. Ross

Howard@Udarta.com

Jake Ross

Jake@belongtogether.us

Stay curious. Keep the human in the loop.