THE CRAFT OF DIALOGUE

Conversing with People Who May Not Be Seeking Truth

Applied Pedagogy Research Lab

Guido Bartolucci, Principal Investigator

guido@appliedpedagogy.com

LAB.APPLIEDPEDAGOGY.COM

SP-001 · April 2026

Research conducted by AI agents (Claude, Anthropic) under human direction.
See LAB.APPLIEDPEDAGOGY.COM for methodology and verification framework.

CONTENTS

1

PART I: CONCEPTUAL MAP

1.1 WHY MOST DIFFICULT CONVERSATIONS FAIL BEFORE THEY BEGIN

You are at a dinner, and someone says something wrong. Not ambiguously wrong — substantively, demonstrably, the-evidence- is-clear wrong. You know the evidence. You could cite it. So you say, carefully, with what you believe is precision and good faith: "Actually, I think the data shows something different." And the conversation immediately goes sideways. Not because your evidence was bad. Not because your tone was hostile. But because you and the other person were playing different games, and neither of you knew it.

You thought you were in an inquiry — two people trying to figure out what's true. They thought they were in a negotiation — two identities competing for the right to define what counts as reasonable. You offered evidence; they heard a status challenge. You asked a clarifying question; they heard cross-examination. You genuinely wanted to understand their position; they genuinely believed you were trying to humiliate them. Neither of you was wrong about what was happening from where you stood. Both of you were wrong about what was happening overall.

This is not a failure of intelligence, and it is not primarily a failure of communication. It is a failure of diagnosis. Most difficult conversations fail not because of bad arguments or insufficient information, but because participants are playing different games and have no map for recognizing that fact. The person attempting inquiry is talking to someone conducting an eristic quarrel. The person trying to deliberate is talking to someone who is only negotiating. The person who wants to establish the facts is talking to someone defending a moral identity. Without a diagnostic vocabulary for what conversational games are possible, you cannot tell which game you are in — and without that diagnosis, every technique in the playbook is a shot in the dark.

This review provides the diagnostic vocabulary. It then provides the playbook. And it provides an honest assessment of when the playbook will not work — when the conversation is not convertible to joint inquiry, and the realistic goal is something narrower: containment, de-escalation, graceful exit, or protecting third parties who are listening.

The conceptual map builds in three layers, stacked from most structural to most evaluative. The first layer, drawn from Douglas Walton and Aristotle, addresses the question: What type of dialogue is this? The second layer, from David Bohm, adds a quality dimension that cuts across all types: is meaning actually flowing, or is it being defended against? The third layer, from Frans van Eemeren's pragma-dialectics, provides a normative backbone: within whatever type of dialogue is operative, is it being conducted in good faith or bad?

Together, these three layers constitute the diagnostic map the practitioner needs before choosing any moves. They also make a single integrative point that the whole review will return to repeatedly: conversation is not a contest to be won. It is a game with constitutive rules, and those rules can be violated — in which case you are no longer playing the game you thought you were playing. The most important diagnostic question is not "Am I making good arguments?" but "Are we both playing by the same rules, and do those rules point toward truth?"

1.2 A NOTE ON SOURCES AND PROVENANCE

This review draws on sources with different levels of verification. Some were read directly by the agent conducting the research. Some were accessed through structured summaries prepared by the PI. Some were verified through abstracts and metadata. A small number — approximately ten percent of claims — rely on training knowledge reconstructed from secondary literature rather than direct access to the primary text.

Three sources in particular require this flag.

Motivational Interviewing. The MI scripts and techniques described throughout this review — the OARS framework, reflective listening levels, the righting reflex, change talk versus sustain talk — are reconstructed from peer-reviewed papers about MI, from clinical training literature, and from the extensive secondary tradition rather than from direct reading of Miller and Rollnick's primary text (3rd ed., 2013). The scripts are approximate rather than verbatim. The underlying principles are well-established and cross-validated by multiple independent reviews (Lundahl et al., 2010; Hettema, Steele, & Miller, 2005), but the specific phrasings used in this review's playbook should be treated as the agent's reconstructions of MI technique, not as direct quotations from Miller and Rollnick.

Stone, Patton, and Heen's *Difficult Conversations*. The three-conversation framework and the contribution frame are similarly reconstructed from secondary sources and the extensive literature that builds on the Harvard Negotiation Project's work.

Van Eemeren's pragma-dialectics. The ten rules for critical discussion are well-established and widely reproduced in the argumentation theory literature, but the specific formulations and diagnostic applications in this review draw on both the primary texts accessed through summaries and on the broader pragma-dialectical tradition.

Where direct access to the primary text was available — Walton's *Methods of Argumentation*, Bohm's *On Dialogue*, Boghossian and Lindsay's *How to Have Impossible Conversations*, and the descriptive foundations (Haidt, Mercier & Sperber, Tavris & Aronson, Kahneman, Aristotle, Perelman) — claims are cited with specific section references and verified against the source text. The provenance status of every source is recorded in 'sources.md'.

Street Epistemology's empirical status. SE as a packaged technique has no peer-reviewed randomized controlled trial evaluation. The supporting evidence is indirect: the non-confrontational questioning approach is supported by MI literature (Lundahl et al., 2010); the focus on epistemological mechanisms over content is formalized by Schmid and Betsch (2019, *PNAS*); narrative elicitation has RCT support via Broockman and Kalla (2016, *Science*). The SE community has produced over 1,000 documented conversations (Magnabosco's corpus) but without control conditions or outcome measurement. When this review recommends SE-derived techniques, the reader should understand that the empirical support comes from adjacent research traditions rather than from direct evaluation of the SE protocol itself.

1.3 LAYER 1 — DIALOGUE TYPE: WHAT GAME IS THIS?

1.3.1 *The Six Games*

Douglas Walton, building on foundational work with E.C.W. Krabbe (1995), identified six primary types of dialogue, each defined by its initial situation (what brought the participants together), its collective goal (what the conversation is for), and its individual participant goals (what each party is trying to achieve). Walton's canonical formulation: "The six basic types of dialogue so far defined

are persuasion dialogue, inquiry, negotiation dialogue, information-seeking dialogue, deliberation and eristic dialogue" (Walton, 2013, §8.1.6, citing Walton & Krabbe, 1995).

The table below summarizes the six types. After it, each type is examined as the practitioner experiences it from the inside.

Dialogue type	Initial situation	Collective goal	What "winning" means
Persuasion	Conflict of opinions	Resolve by rational argument	Prove your thesis; other party updates
Inquiry	Need to establish truth	Prove or disprove a proposition	Truth is established; everyone updates
Deliberation	Need to decide on action	Decide the best course	Best decision reached for all
Negotiation	Incompatible interests	Mutually acceptable settlement	Best deal for self within settlement
Information-seeking	One party lacks knowledge	Transfer information	Info transferred accurately
Eristic (quarrel)	Personal antagonism	Express feelings, reveal conflict	Defeat the opponent personally

The table's most important column is the last one. What "winning" means is the key to diagnosing which game is being played — because different players in the same conversation can have radically different notions of what success looks like.

1.3.2 *Persuasion Dialogue*

Persuasion dialogue begins with a genuine conflict of opinions. One party holds a thesis that another contests or denies. The collective goal is to resolve that conflict through rational means — by pitting the strongest arguments on each side against each other until one side's position is demonstrably better supported.

From the inside, persuasion dialogue feels like intellectual engagement: you are making your best case, tracking the other person's objections, offering evidence when challenged. There is an adversarial element — you want to prove your thesis — but that adversarial element is constrained by constitutive rules. A rational arguer, in Walton's formulation, "is one who follows the protocols for the type of dialogue appropriate for the argumentation in which he is engaged" (§2.4.6). That means: using only premises the respondent has accepted, opening yourself to critical questioning, and — crucially — being willing to retract when your argument fails.

Diagnostic signals you are in persuasion dialogue: Both parties are tracking each other's arguments. Evidence is offered and evaluated on its merits. Critical questions are welcomed rather than deflected. When one party makes a point the other cannot immediately answer, there is a pause — a genuine moment of consideration rather than an immediate pivot to a new attack.

What "winning" means: Discharging the burden of proof at the standard set when the dispute opened. In genuinely good-faith persuasion dialogue, winning and losing are not primarily social events — they are epistemic ones. The person who was wrong updates. The relationship often improves, because both parties have demonstrated that they can handle the truth.

Rules of good-faith participation: Use only premises the other party has accepted or would accept. Defend your thesis when challenged. Attend to the actual argument made, not a distorted version of it. Be prepared to retract your position if it is successfully refuted.

Transition affordances: Persuasion dialogue regularly and legitimately pauses for information-seeking — one party requests evidence that the other must retrieve. It can also legitimately embed into inquiry when a disputed factual claim needs independent verification before the argument can continue. The danger zone is the transition to eristic: when the adversarial element overwhelms the cooperative element, persuasion deteriorates into quarrel.

1.3.3 *Inquiry Dialogue*

Inquiry dialogue begins not with a conflict between parties but with a need to establish whether a proposition is true or false — a need that neither party can yet meet alone. Unlike persuasion, inquiry has no protagonist defending a predetermined thesis. All participants work toward the same goal: establishing the truth of a designated statement.

From the inside, this is the rarest and most cognitively satisfying form of conversation. Both parties genuinely do not know the answer. Error discovered by either party is gain for both. The question drives the conversation, not the parties' prior commitments.

Inquiry dialogue sets the highest standard of proof — what Walton calls cumulativeness: "once a statement is accepted in an inquiry, it cannot be retracted without reopening the inquiry" (§8.9.1). This is why genuine scientific inquiry is so demanding. You cannot just say something and move on — every accepted premise becomes part of the foundation on which subsequent reasoning builds.

Diagnostic signals you are in inquiry: Neither party entered with a predetermined conclusion. Both can articulate what evidence would change their minds. When one party discovers an error in their own reasoning, they announce it rather than hoping it goes unnoticed. The conversation moves forward cumulatively rather than cycling through the same ground.

What "winning" means: There is no winner in inquiry. Truth is established; everyone updates. This is exactly the game the review's target reader wants to be playing and rarely finds themselves in when conversations become difficult.

Transition affordances: Inquiry often follows what Walton calls discovery dialogue — the exploratory phase where hypotheses are generated before the standards for proving them are set. It legitimately embeds information-seeking whenever evidence needs to be gathered.

1.3.4 *Deliberation Dialogue*

Deliberation begins with a need to decide what action to take. The question is not whether a proposition is true but what to do — and both parties have a stake in the outcome. The collective goal is the best available course of action for the situation.

Deliberation is the appropriate dialogue type for most practical decisions: where to live, how to allocate organizational resources, what policy to adopt. From the inside, it feels cooperative in structure but potentially contested in content. Both parties want a good decision; they may disagree about what counts as good.

A key feature of deliberation: no one enters with a burden of persuasion to discharge. Proposals can emerge from anywhere, including from the party that initially opposed a direction. A deliberator who ends up supporting a proposal they initially opposed has not "lost" — they have contributed to a better decision.

Diagnostic signals you are in deliberation: The question being debated is practical — "what should we do?" rather than "who is right?" Arguments from consequences ("if we do X, Y will happen") and from values ("what matters most to us here?") dominate. Both parties can imagine being persuaded by a good argument from the other side.

What "winning" means: The best decision is reached, given what was known. This means there is no individual winner — though individual parties may take credit for the winning proposal. The danger: parties slide from "best decision for all" to "decision that vindicates my prior position." When that happens, deliberation has begun to drift toward negotiation.

Transition affordances: Deliberation legitimately embeds information-seeking (when facts are needed to evaluate options) and can legitimately shift into negotiation (when parties need to divide resources required to execute the agreed plan). The shift from deliberation to eristic is the most disorienting: the conversation was about what to do, and suddenly it is about who is right or who to blame.

1.3.5 *Negotiation Dialogue*

Negotiation begins when parties have incompatible interests and need a settlement. Unlike deliberation, where the goal is the best decision for all, negotiation explicitly acknowledges that parties have different interests — and that the goal is to find a settlement that is mutually acceptable, even if not optimal for any individual party.

From the inside, negotiation is openly strategic. Arguments are evaluated not primarily for their truth-value but for their power to move the other party. "Will this offer shift them?" is the operative question, not "Is this premise true?" The Harvard Negotiation Project's principled negotiation — focusing on interests rather than positions, using objective criteria, separating people from problems — represents an attempt to move negotiation toward something more like deliberation without denying its fundamentally interest-based structure (Fisher & Ury, 2011).

Diagnostic signals you are in negotiation: Offers and counteroffers are being made, not just arguments. Each party has a clear sense of what they need and what they can concede. The conversation tracks a zone of potential agreement rather than a verdict on who is correct. Parties are conscious of their best alternative to a negotiated agreement.

What "winning" means: The best deal for yourself within the terms of a settlement the other party will actually accept. This is the most explicitly self-interested goal among the six types — and that is not a criticism. Legitimate interests exist; negotiation is the appropriate mechanism for reconciling them when they conflict.

Transition affordances: Negotiation legitimately embeds deliberation when the parties need to agree on criteria for evaluating offers. The canonical illicit shift is from negotiation to eristic. Walton's example is instructive: "a union– management negotiation deteriorates into an eristic dialogue in which each side bitterly attacks the other in an antagonistic manner. This kind of shift is not an embedding, because quarreling is not only unhelpful to the conduct of the negotiation but is antithetical to it, and may very well even block it altogether, by leading to a strike, for example" (§8.8.2). The quarrel does not serve the negotiation's collective goal — it destroys it.

1.3.6 *Information-Seeking Dialogue*

Information-seeking begins when one party lacks information that another possesses. The collective goal is simple: transfer the relevant information accurately and completely.

This is the most structurally asymmetric of the six types — one party knows something the other does not — and that asymmetry can generate problems. From the inside, a genuine information-seeking conversation feels cooperative and relatively low-stakes: you ask questions, the other party answers, both parties are satisfied when the information has moved. The failure mode is when the asymmetry becomes a power relationship — when the possessor of information uses their position to manipulate rather than inform.

Diagnostic signals you are in information-seeking: One party is asking questions; the other is answering. The questioner is genuinely interested in accurate information, not in obtaining confirmation of a predetermined view. The answerer is motivated to transfer information accurately.

What "winning" means: The information has moved. Everyone has what they need to proceed. Winning in information-seeking is essentially invisible — it just means the conversation accomplished its purpose.

Transition affordances: Information-seeking is the most commonly embedded dialogue type. It appears inside persuasion (when one party requests evidence), inside deliberation (when options need to be evaluated against facts), and inside inquiry (when background knowledge is required). The danger: information-seeking disguised as inquiry — when the "question" is actually a demand for confirmation of what the asker already believes.

1.3.7 *Eristic Dialogue (The Quarrel)*

Eristic dialogue — from the Greek *eris*, strife — begins with personal antagonism or conflict between parties. There is no constructive collective goal. The purpose of an eristic exchange, at best, is to reveal the depth and nature of the conflict. At worst, it simply escalates it.

Eristic dialogue is the central category for this review's purposes. Not because we want to play it, but because recognizing when we have been pulled into it — or when the other party is already there — is the most important diagnostic skill a practitioner can develop.

From the inside, eristic dialogue does not feel like a quarrel. It feels like being profoundly right. The other person is obviously wrong, obviously unfair, obviously operating in bad faith. The desire to prove this, to expose it, to make them admit it — that desire is the signature of eristic mode. Walton notes that "bad things can happen, including the committing of fallacies by shifting to a quarrel that is purely adversarial. Any conversation in which reasonable argumentation is to be used for some constructive purpose must strike the right balance between this adversarial aspect and the need for all parties to follow the Gricean maxims of polite conversation" (§9.1.14).

Diagnostic signals you are in eristic dialogue: The other party is not tracking your arguments — they are waiting for you to finish so they can make their next attack. Emotional intensity is high and disproportionate to what is actually at stake. Each new argument from either party is met with escalation, not consideration. The subject has shifted from the original topic to the character and motives of the parties. "Winning" now means making the other person look bad, not establishing what is true.

What "winning" means: Defeating the other person personally. Making them capitulate, look foolish, or lose the audience's esteem. This is categorically different from every other form of winning — it has no epistemic content whatsoever. A better argument does not win an eristic exchange; only dominance does. This is why trying to win an eristic exchange through superior reasoning is a category error.

Rules of eristic — or rather, the absence of rules: Eristic dialogue has no constitutive rule that rewards truth-tracking. In persuasion, the rules reward whoever produces the stronger argument. In eristic, the rules — such as they are — reward whoever defeats the other person, by whatever

means. Ad hominem, personal attack, emotional manipulation, and status games are characteristic moves in eristic precisely because they work within eristic's terms even when they are completely irrelevant to the substantive question.

Transition affordances: Eristic dialogue can shift back toward persuasion, but only when both parties explicitly accept a new set of constitutive rules. One party can attempt this unilaterally — by stopping, naming what is happening, and offering to return to the argument — but the other party must accept the invitation. The moves for attempting this transition are in Part II.

1.3.8 *The Dialectical Shift: When Games Change*

Every practitioner needs to understand not just the six types but how conversations move between them. Walton's concept of the **dialectical shift** refers to transitions between dialogue types during the course of an exchange. "A dialectical shift is said to occur in cases where, during a sequence of argumentation, the participants begin to engage in a type of dialogue different from the one in which they were initially engaged" (§8.8.2, citing Walton & Krabbe, 1995).

Not all shifts are problematic. Many are not only legitimate but necessary:

- A persuasion dialogue legitimately pauses to information-seek when one party needs to retrieve evidence.
- A deliberation legitimately embeds into negotiation when parties need to divide resources required to execute a decision.
- A discovery process legitimately transitions into inquiry when a hypothesis is strong enough to warrant formal evaluation.

These are **embeddings** — the second dialogue type serves the first. "In this kind of case, we say that one dialogue is embedded in the other, meaning that the second dialogue fits into the first and helps it along toward achieving its collective goal" (§8.8.2).

The **illicit shift** is different in kind, not just degree. An illicit shift occurs when the second dialogue type does not serve the first but actively undermines or destroys it. The union-management example is canonical: the eristic quarrel that emerges from a failed negotiation does not help the negotiation — it makes the negotiation impossible. The parties are no longer playing the game; they are fighting.

Illicit shifts are the primary mechanism through which conversations fail. The most important illicit shift for this review is the deterioration of persuasion dialogue (or any constructive type) into eristic. When this happens, the practitioner faces a choice: attempt to recover the original dialogue type by naming the shift and proposing a return to the rules, or recognize that the original game is no longer available and shift goals accordingly.

The masking of illicit shifts is what makes them so dangerous. Walton observes that bad-faith participants often perform the surface forms of persuasion dialogue — they use premises, draw conclusions, cite evidence — while actually playing eristic. They have the appearance of argument without the constitutive commitments that argument requires. This is the distinction between a **paralogism** and a **sophism**.

A paralogism is an honest reasoning error. The arguer has made a blunder — produced a formally defective argument, misapplied an argument scheme, drawn a conclusion that does not follow from their premises. They are still trying to argue correctly. The appropriate response is correction: "I think there's a gap in that argument — here's where it breaks down."

A sophism is a deliberate tactic. "It is based on the idea that an organized rule-governed dialogue in which arguments are exchanged, like a critical discussion, is partly adversarial but also partly

cooperative" (§9.1.8). The sophistical arguer is not reasoning badly — they are using the forms of reasoning to achieve the goals of eristic: to defeat, humiliate, or silence the other party while appearing to engage on the merits. The appropriate response is not correction but game-recognition: the conversation has shifted, and the practitioner must decide what to do about that.

The distinction between paralogism and sophism does not require imputing motives. In practice, you often cannot know whether an error is honest or tactical. The practical test is behavioral: does the person update when the error is pointed out? A paralogist corrects course; a sophist pivots to a new tactic. The response to paralogism is corrective; the response to sophism is diagnostic — naming the pattern and reconsidering what game is now being played.

1.3.9 *Aristotle's Three Species of Rhetoric: A Cross-Cutting Dimension*

Walton's typology addresses the structure of dialogue — what participants are trying to accomplish together. Aristotle's three species of rhetoric, from the *Art of Rhetoric* (Bartlett trans.), address the temporal orientation of the dispute. They cross-cut Walton's types in a way that is practically important.

The **deliberative** species (*symbouleutikon*) concerns the future: What should we do? Its paradigmatic context is the assembly deciding on policy. Arguments are about benefits, harms, and the consequences of different courses of action. The dominant questions are practical: what will work, what is best, what do we value enough to pursue at cost?

The **forensic** species (*dikanikon*) concerns the past: Who is responsible? What happened? Its paradigmatic context is the law court. Arguments are about evidence, causation, and attribution. The dominant questions are factual and moral: Did this happen? Who did it? Was it justified?

The **epideictic** species (*epideiktikon*) concerns praise and blame: Who is praiseworthy? Who is blameworthy? Its paradigmatic context is the ceremonial occasion — the funeral oration, the celebratory speech. Arguments are about character, virtue, and worth. The dominant questions are evaluative: Is this person good? Do they deserve honor or censure?

These three species map onto different conversational registers that appear across all of Walton's dialogue types. A deliberation can be conducted in epideictic mode — instead of asking "what should we do?" both parties are implicitly arguing "my position is the virtuous one." A negotiation can veer into forensic mode — instead of seeking a settlement, both parties are reconstructing who wronged whom. An eristic exchange is almost always epideictic in register — the parties are primarily establishing who is a good and who is a bad person.

The practical utility of Aristotle's three species is that they provide a fast diagnostic for getting stuck. When a conversation that started as deliberation has become stuck, the first question to ask is: has the register shifted from deliberative to epideictic? Are we now arguing about who is right rather than what to do? The temporal question — future-oriented versus past-oriented versus character-oriented — is often easier to detect in the moment than the more structural question of which Walton type is operative.

When a deliberative conversation has gone epideictic, the recovery move is to return to the future: "Let's set aside who made which decision and talk about what to do from here." When a persuasion dialogue has gone forensic — arguing about whose past behavior was justified — the recovery move is often to disentangle the factual question from the responsibility question and address them sequentially. Part II of this review develops these recovery moves as named techniques.

1.4 LAYER 2 — DIALOGUE QUALITY: DISCUSSION OR DIALOGUE?

Walton's typology tells us what kind of conversation we are in. David Bohm's distinction tells us how deeply we are actually engaging. This is a quality dimension, not a sixth type — it applies across all of Walton's categories and describes how much meaning is actually flowing between participants.

Bohm's derivation begins with etymology. "Dialogue" comes from the Greek *dialogos*: *dia* meaning "through" (not "two") and *logos* meaning "the word" or "meaning." "The picture or image that this derivation suggests is of a stream of meaning flowing among and through us and between us" (Bohm, 1996, §6.0.2). "Discussion," by contrast, "has the same root as 'percussion' and 'concussion.' It really means to break things up" (§6.0.3). Discussion is analysis: competing viewpoints, tactical exchange of positions, everyone presenting a different view and defending it against attack.

Bohm's characterization is blunt: "Discussion is almost like a ping-pong game, where people are batting the ideas back and forth and the object of the game is to win or to get points for yourself" (§6.0.3). In genuine dialogue — his dialogue, the aspirational pole — "there is no attempt to gain points, or to make your particular view prevail. Rather, whenever any mistake is discovered on the part of anybody, everybody gains. It's a situation called win-win, whereas the other game is win-lose" (§6.0.4).

The critical structural difference: in discussion mode, finding out you are wrong means you have lost. In genuine dialogue mode, finding out you are wrong — or finding out that the other person is wrong, or finding out that both parties were working with an incomplete picture — is a gain for everyone. The error was a cost; its discovery is a benefit. When that shift in orientation has occurred, the conversation's entire character changes.

What Bohm calls "discussion" maps roughly onto Walton's persuasion and eristic types in their more defended forms. What he calls "dialogue" corresponds most closely to Walton's inquiry type conducted in ideal conditions — but Bohm's account adds something that Walton's formal typology does not reach: a phenomenological description of what is happening inside participants when genuine meaning-sharing occurs.

Most of what passes for dialogue in institutions, Bohm observes, stalls at what he calls the negotiation stage — "a very preliminary stage of dialogue" in which participants find a common way to proceed without ever touching their fundamental assumptions (§6.0.41). Trade-offs are reached. Agreements are signed. But the deeper conflict — the one rooted in assumptions that neither party has examined — remains untouched and will resurface.

1.4.1 *Why Genuine Dialogue Is Rare: Opinions as Identity*

The structural obstacle to genuine dialogue is not stupidity or bad faith. It is the relationship between opinions and identity. Bohm's analysis: "It is important to see that the different opinions that you have are the result of past thought: all your experiences, what other people have said, and what not. That is all programmed into your memory. You may then identify with those opinions and react to defend them. But it doesn't make sense to do this. If the opinion is right, it doesn't need such a reaction. And if it is wrong, why should you defend it? If you are identified with it, however, you do defend it. It is as if you yourself are under attack when your opinion is challenged" (§6.0.11).

The mechanism Bohm describes has deep roots in what he calls the transference of self-defense: "The natural self-defense impulse, which we got in the jungle, has been transferred from the jungle

animals to these opinions. In other words, we say that there are some dangerous opinions out there — just as there might be dangerous tigers. And there are some very precious animals inside us that have to be defended" (§6.0.97). The impulse is not irrational within its own logic — the animal protecting its territory is doing exactly what it evolved to do. The problem is that the territory is now intellectual, and the appropriate response to a challenging argument is examination, not defense.

This analysis is directly corroborated by Mercier and Sperber's argumentative theory of reason (2011). Their central claim: reasoning evolved not primarily for private truth-seeking but for the social function of producing and evaluating arguments in a communicative context. Confirmation bias — the tendency to seek evidence supporting one's existing position — is not a design flaw in human cognition. It is a feature of a system adapted to convince, not to find truth. "The confirmation bias contributes to an efficient form of division of cognitive labor" (Mercier & Sperber, 2011, p. 65): each party makes the strongest case for their position; the group then filters through argument evaluation. The whole system works for the group when there is genuine disagreement and a shared interest in truth. It works against the individual when they are reasoning alone.

The practical implication is stark: most difficult conversations fail not because participants are reasoning badly but because they are reasoning exactly as their cognitive architecture intends. They are defending positions, seeking confirming evidence, producing arguments — all in service of the social goal of maintaining their standing and protecting their identity. Shifting from that mode to genuine inquiry is a real achievement, not the default.

1.4.2 *Suspension: The Practice That Makes Dialogue Possible*

Bohm's most practically useful concept is **suspension**. To move from discussion to dialogue mode — from defending opinions to genuinely examining them — participants must be able to suspend their assumptions. Suspension does not mean abandoning assumptions; it means holding them at arm's length so they can be observed.

Bohm's formulation is precise: "We have been saying that people in any group will bring to it their assumptions, and as the group continues meeting, those assumptions will come up. Then what is called for is to suspend those assumptions, so that you neither carry them out nor suppress them. You don't believe them, nor do you disbelieve them; you don't judge them as good or bad" (§6.0.48). The key three-part exclusion: suspension is not acting on the impulse, not suppressing it, and not judging it.

Bohm's illustration using anger is worth quoting at length because it clarifies what suspension is not: "Normally when you are angry you start to react outwardly, and you may just say something nasty. Now suppose I try to suspend that reaction. Not only will I now not insult that person outwardly, but I will suspend the insult that I make inside of me. Even if I don't insult somebody outwardly, I am insulting him inside. So I will suspend that, too. I hold it back, I reflect it back. You may also think of it as suspended in front of you so that you can look at it — sort of reflected back as if you were in front of a mirror" (§6.0.48).

The mirror image captures what suspension enables: you can see the impulse, its structure, its energy, the thought underneath it — from a position outside the automatic reaction. What remains invisible when you suppress ("I'm not angry") becomes visible when you suspend.

Suppression, Bohm notes with precision, is not suspension: "If a person is aggressive... At some point he may notice what is happening. He may think, 'I am aggressive, I mustn't be aggressive.' That suppresses the action, which means that you are still aggressive, against yourself. So it hasn't changed. The observer of aggression is pervaded with aggression" (§10.0.2). The person who

suppresses their defensiveness while remaining defensive is not suspending — they are redirecting the same energy inward.

Why does this matter for conversation practice? Because the alternative to genuine suspension is performed openness — appearing to consider the other person's point while actually mobilizing defenses. Performed openness is detectable. The other person, whose own emotional sensitivity is calibrated for exactly this, will notice that the words say "interesting point" while the posture says "threat." The conversation will not shift into genuine dialogue because the conditions for genuine dialogue — both parties actually looking at the territory rather than defending their positions — are not present.

1.4.3 *Proprioception of Thought*

Bohm extends the suspension concept through his most distinctive metaphor: proprioception of thought. The body has a proprioceptive sense — it knows where its limbs are, can feel itself moving, can distinguish its own actions from external forces acting upon it. Bohm proposes that the mind can develop something analogous: the capacity to perceive its own thinking as it happens, including the gap between an intention and the thought's result.

The illustration that makes this vivid: a woman who woke in the night after a stroke. Her motor nerves worked but her sensory nerves did not. She was hitting herself. Because she could not feel that she was the source of the contact, she assumed someone else was attacking her — and the more she defended herself, the worse the "attack" became. "The proprioception had broken down. She no longer saw the relation between the intention to move and the result" (§6.0.63). The light being turned on established a new proprioceptive loop through sight; she could see that she was the source.

Thought, Bohm argues, operates similarly when proprioception is absent. An emotional reaction is produced by a thought; the thought is not tracked as thought; the emotional reaction is therefore experienced as a direct perception of external reality. The feeling of being wronged, the sense of being under attack, the conviction that the other person is obviously in bad faith — these can all be, at least in part, products of thought that thought itself has generated and then lost track of. "Practically all the problems of the human race are due to the fact that thought is not proprioceptive" (§6.0.66).

The connection between suspension and proprioception is direct. "The point of suspension is to help make proprioception possible, to create a mirror so that you can see the results of your thought" (§6.0.65). Suspension is the practice; proprioception is what the practice makes possible. When you pause a defensive impulse and hold it up for examination, you can see — sometimes with considerable surprise — that the "attack" you were about to respond to was a reaction to your own thought about what was said, not necessarily to what was said.

This metacognitive capacity — the ability to notice, in real time, when your own thinking has shifted from examination to defense — is what makes genuine dialogue possible at the individual level. It is also, Bohm suggests, what is most systematically underdeveloped in ordinary cognitive life. The layer of judgment and metacognition that this review aims to cultivate is, in Bohm's terms, the development of proprioception of thought.

1.5 LAYER 3 — NORMATIVE ASSESSMENT: GOOD FAITH OR BAD FAITH?

Walton's typology tells us what game is being played. Bohm's quality dimension tells us how deeply meaning is flowing. But neither tells us whether a participant is playing by the rules of the game

they claim to be in. That is the function of van Eemeren's pragma-dialectics — the third and most evaluative diagnostic layer.

Van Eemeren and Grootendorst's project began in the 1970s and reached its fullest expression in their 1992 *Argumentation, Communication, and Fallacies* and the 2004 *A Systematic Theory of Argumentation*. The aim is explicitly normative: not to describe how people argue but to specify how they ought to argue "if their goal is resolving a difference of opinion through rational means" (van Eemeren & Grootendorst, 1992). They identify ten rules for the **critical discussion** — their idealized version of persuasion dialogue — and define a fallacy as any move that violates one of those rules.

For the practitioner, the value of this framework is the negative. The ten rules are most useful not as a checklist for conducting ideal discussions but as a **bad-faith diagnostic**: each rule, when violated, marks a specific and nameable form of bad-faith move. Knowing the rules means recognizing their violations in real time.

1.5.1 *The Ten Rules as a Diagnostic Instrument*

Rule 1 — Freedom: Neither party may prevent the other from advancing or questioning standpoints. Violation: "You're not an expert, so you're not entitled to say anything about this at all" (Walton, §9.1.12). This move — delegitimizing the person rather than their argument — is the clearest possible signal that the conversation has shifted toward eristic. When the target of attack is the speaker rather than the claim, the constitutive rules of persuasion dialogue have been violated.

Rule 2 — Burden of proof: The party who advances a standpoint must defend it when asked. Violation: Demanding that the other party disprove your claim without offering any positive account of why it should be accepted. The Street Epistemology question "What would it take to change your mind?" is a direct application of this rule — it makes the burden-of-proof structure visible and asks the interlocutor to own it.

Rule 3 — Standpoint: Attacks must address the standpoint actually advanced, not a distorted version. Violation: The straw man. Someone who misrepresents your position — through distortion, exaggeration, selective quotation, or attribution of premises you did not advance — is violating Rule 3. This is one of the most prevalent bad-faith moves in public argument and one of the hardest to name gracefully in the moment.

Rule 4 — Relevance: Arguments must be relevant to the standpoint in question. Violation: Ad hominem attacks that function as substitutes for argument; appeals to emotion that bypass rather than supplement substantive reasoning; appeals to popularity ("everyone knows") as evidence for truth. The rule does not ban emotional appeals entirely — it bans them when they replace rather than accompany substantive argument.

Rule 5 — Unexpressed premises: Unexpressed premises may only be attributed if they would actually be endorsed. Violation: Uncharitable inference — attributing to the interlocutor a premise more extreme or objectionable than they would own. This rule is central to the "narrative-pushing" hard case examined in Part III: someone who holds a vague intuition they cannot fully articulate is vulnerable to having implicit premises falsely attributed, which then get attacked rather than the actual position examined.

Rule 6 — Starting points: Premises must not be presented as shared if they are not. Violation: Begging the question — treating a disputed proposition as an accepted starting point. Identifiable by the construction "Of course X is true, which is why Y follows" — the "of course" embeds a contested claim as given. Also includes false dichotomy: presenting two options as exhaustive when the alternative space is larger.

Rule 7 — Argument scheme: Argument schemes must be applied correctly. Violation: Misapplied analogies (treating superficially similar cases as relevantly similar); slippery slope without a specified causal chain; false causation from correlation. A slippery slope argument is legitimate when the recursive causal structure is actually specified; it becomes fallacious when the chain is merely asserted.

Rule 8 — Validity: Conclusions must follow from premises. Violation: Formal logical fallacies — affirming the consequent, denying the antecedent. Less common in ordinary conversation than Rules 1-6 violations, but structurally important: when a conclusion is presented as certain when it merely follows from premises that have not been examined, the implicit step needs surfacing.

Rule 9 — Closure: Failed defenses require the protagonist to retract the standpoint; successful defenses require the antagonist to withdraw doubt. Violation: Continuing to advance a position after it has been successfully refuted. Refusing to concede when the argument demands concession. Claiming victory when the argument has not been successfully concluded.

Rule 9 is the most diagnostic single rule for this review's central question. It is the formal specification of what "updating when shown evidence" requires. A participant who violates Rule 9 — who maintains their position regardless of what the argument has established — is not engaged in a critical discussion. They are engaged in something else, and the practitioner needs to recognize that and respond accordingly.

Rule 10 — Language: Language must be clear and unambiguous; the other party's formulations must be interpreted as carefully and accurately as possible. Violation: Strategic vagueness — exploiting ambiguity so that the claim can be understood in whatever way is most convenient at any moment. Shifting definitions mid-argument. The positive requirement — interpret the other party's formulations charitably — is the pragma-dialectical analogue of the principle of charity.

1.5.2 *Schopenhauer's Catalog: The Practitioner View from the Inside*

Schopenhauer's *The Art of Always Being Right* (posthumously published from notes; English translation widely available) presents the same bad-faith moves from the practitioner's perspective — how to deploy them, not just recognize them. The book lists 38 stratagems for winning arguments regardless of whether you are right, including: extending the opponent's thesis beyond its intended meaning so it becomes absurd (violation of Rules 3 and 5); appealing to authority in place of argument (Rule 4); making the opponent angry so that they argue badly (eristic tactic within persuasion shell); and claiming victory in the face of counterargument simply by asserting it (Rule 9).

The review's treatment of Schopenhauer is diagnostic, not prescriptive. The 38 stratagems are a detection instrument: knowing how these moves are deployed means recognizing them when they are deployed against you. Each stratagem is a specific form of illicit shift — persuasion dialogue's constitutive rules being violated while its surface forms are preserved. The appropriate response is game-recognition followed by the decision: do I attempt to recover the original dialogue type, or have conditions deteriorated to the point where a different goal is more realistic?

1.6 THE COMPOSITE DIAGNOSTIC: WHEN CAN THIS CONVERT TO JOINT INQUIRY?

Note: The following diagnostic is an agent construction — a synthesis of frameworks from multiple source traditions integrated into a single instrument. It is not attributable to any single source and should be used as a working tool, not as established theory.

After presenting the three diagnostic layers — dialogue type (Walton), dialogue quality (Bohm), and normative evaluation (van Eemeren) — the practitioner needs a fast, practical test. Not a comprehensive analysis but a triage: is conversion to genuine joint inquiry achievable, or is the realistic goal narrower?

Five tests, each derived from a different source tradition. Apply all five. When two or more light red, conversion to genuine joint inquiry is unlikely in the current exchange. The realistic goals then narrow to: containment (Eddy's BIFF method), de-escalation (Thompson's Verbal Judo), graceful exit (Boghossian & Lindsay's disengagement section), or protecting third parties who are listening (Rauch's point that the audience often matters more than the interlocutor).

Test 1 — Disconfirmability (from Boghossian & Lindsay): Ask, directly or indirectly: "Under what conditions could that belief be wrong?" Boghossian and Lindsay describe this as "the single most effective technique to instill doubt and help people change their minds" (§0.60.1). If the person cannot generate any conditions under which their belief would be false — if every proposed scenario is deflected or explained away — the belief is not held on evidential grounds. This does not necessarily mean the person is arguing in bad faith; it means the epistemic structure of their commitment is unfalsifiable, which has direct implications for what conversational moves can accomplish.

Test 2 — Emotional proportionality (from Eddy): Are emotional reactions dramatically disproportionate to what the topic warrants? A vigorous argument about a policy question where the stakes are genuinely high warrants emotional engagement. The same vigorous argument about a claim that scarcely affects either party's life, conducted with the intensity of a personal attack — that disproportion is a diagnostic signal. Bill Eddy's WEB (Words, Emotions, Behaviors) method for identifying high-conflict personality dynamics emphasizes this proportionality question. Emotional reaction calibrated to actual stakes suggests a person who can be reasoned with; emotional reaction calibrated to ego threat suggests that conversion to joint inquiry may require reducing the perceived threat first, not advancing the argument.

Test 3 — Argument tracking (from Walton's eristic diagnostic): Is the person tracking your arguments at all, or only pushing their own regardless of what you say? In persuasion dialogue, arguments from one party constrain and respond to arguments from the other. A response that does not engage the preceding argument — that simply continues the previous line of attack as if nothing has been said — is a diagnostic signal for eristic mode. The constitutive rules of argument are not in play. The person is not processing your arguments; they are waiting for their turn to push.

Test 4 — Uncertainty acknowledgment (from Galef): Does the person acknowledge uncertainty anywhere, on any topic? The soldier mindset, as Galef analyzes it, treats uncertainty as a tactical liability — something to be concealed rather than reported. Scout mindset treats uncertainty as information — something to be measured and communicated. A person who acknowledges no uncertainty, on any claim, across the full range of a conversation, is displaying the diagnostic signature of what Galef calls the soldier mindset. This is not a moral failing; it is a description of a cognitive mode. But it is information about what conversational moves are likely to succeed. Techniques that invite the person to acknowledge gradations of certainty (the calibration question sequence from Street Epistemology) are precisely designed for this case.

Test 5 — Contestability (from Rauch): Is the person's epistemic regime fallibilist or unfalsifiable? Jonathan Rauch's *The Constitution of Knowledge* describes the epistemic norms that make science, journalism, and liberal democracy possible: the commitment to the principle that any claim can in principle be shown to be wrong, and that this fallibilism is a feature, not a bug. An interlocutor whose epistemic regime is unfalsifiable — who holds beliefs independently of any evidence that could be offered for or against them — is not playing by the rules of any constructive dialogue

type. The appropriate prescription is not better arguments but either institutional defense (making sure the rules of inquiry are maintained in the broader conversation) or graceful exit.

When two or more of these tests light red, the practitioner's calculus shifts. Joint inquiry is not impossible — humans change their minds in unexpected ways, often later than the conversation that prompted it — but it is not achievable as the primary goal of this exchange. The goals that remain available: maintaining your own epistemic integrity; modeling good-faith argument for any third parties present; de-escalating to prevent relationship damage; finding partial agreement on narrower questions; and disengaging without confirming the interlocutor's narrative that the opposition refuses to engage.

The diagnostic is not a license to give up. It is a license to be realistic about what this conversation can accomplish and to choose techniques appropriate to that realistic assessment rather than techniques appropriate to a conversation that is not occurring.

1.7 THE DISPOSITION SECTION: THREE POSTURES THAT MAKE THE MAP USABLE

Knowing the three diagnostic layers and the composite diagnostic is necessary but not sufficient. The practitioner also needs to bring a specific inner orientation to the map — a set of postures that determine whether the conceptual vocabulary produces genuine diagnosis or serves as sophisticated ammunition for the eristic game. This section describes three such postures: scout mindset (Galef), humble inquiry (Schein), and conversational receptiveness (Minson & Yeomans). All three converge on the same underlying requirement, which is the review's first integrative finding: non-judgmental listening is the universal precondition for any of the playbook's techniques to function.

1.7.1 *Scout Mindset (Galef)*

Julia Galef's central distinction, developed in *The Scout Mindset* (2021), is motivational rather than cognitive: the difference between the **soldier mindset** and the **scout mindset** is not in intelligence or knowledge but in what the person is motivated to do with those tools.

The soldier asks: "Can I believe this?" — searching for reasons to accept what is welcome — or "Must I believe this?" — searching for reasons to reject what is unwelcome. The scout asks: "Is it true?" Galef's gloss: "Scout mindset: the motivation to see things as they are, not as you wish they were" (§1.0.2).

The soldier is not being dishonest. From the inside, motivated reasoning feels like objective evaluation. "The tricky thing about motivated reasoning is that even though it's easy to spot in other people, it doesn't feel like motivated reasoning from the inside. When we reason, it feels like we're being objective. Fair-minded. Dispassionately evaluating the facts" (§3.4.1). This is why the soldier/scout distinction is not primarily a distinction between bad people and good people. Anyone can be in soldier mode; everyone is, in some domains and some moments.

The counter-intuitive finding that grounds Galef's framework: scientific intelligence and knowledge amplify ideological polarization on contested empirical questions, not reduce it. Kahan's research (2013) shows that more analytically capable individuals are more polarized on questions like climate change and gun control — because greater analytical capacity means a better ability to rationalize any position, including a deeply defended one. Intelligence is a tool. The soldier uses it for defense; the scout uses it for navigation. "There's nothing inherent to the tools that makes you a scout" (§5.4.18).

The scout mindset is identified behaviorally, not self-perceptually: "The test of scout mindset isn't whether you see yourself as the kind of person who does these things. It's whether you can point to concrete cases in which you did, in fact, do these things" (§5.5.7). Galef's five behavioral markers — telling others when they were right, seeking disconfirming evidence for your own hypotheses, taking precautions against self-deception, having critics you respect — provide a self-assessment the practitioner can run before entering a difficult conversation.

The connection to the playbook is direct: the playbook's techniques are designed to seek truth in difficult conversations. If the practitioner is using them in soldier mode — deploying the forms of genuine inquiry to win rather than to find out — the techniques become manipulation. The scout mindset is not an optional embellishment; it is the condition under which the playbook can accomplish what it is designed to accomplish.

Galef's self-diagnostic: "How would I feel if I were wrong?" If the answer is relief — because you would want to know — you are operating in scout mode. If the answer is dread, distress, or defensive mobilization, the soldier has taken over. This check can be run at any point in a difficult conversation. The answer will often be honest.

1.7.2 *Humble Inquiry (Schein)*

Edgar Schein's *Humble Inquiry* (2nd ed., with Peter A. Schein, 2021) identifies the dispositional prerequisite for questions to function as genuine inquiry rather than as disguised telling. The definition: "Humble Inquiry is the fine art of drawing someone out, of asking questions to which you do not already know the answer, of building a relationship based on curiosity and interest in another person" (§0.0.104).

Three elements are simultaneous in this definition: the epistemic condition (genuinely not knowing), the relational goal (building a relationship), and the motivational source (curiosity and interest). The absence of any one transforms humble inquiry into something else: diagnostic inquiry (steering toward a predetermined conclusion), confrontive inquiry (embedding your own answer in the question's form), or performed curiosity (going through the motions while actually waiting for your turn to tell).

Schein distinguishes four modes of inquiry along a spectrum from genuinely other-oriented to genuinely self-serving:

- **Humble Inquiry**: Pure curiosity, no agenda. Staying on your own side of the net as a listener.
- **Diagnostic Inquiry**: Exploring reactions, feelings, causes — still agenda-neutral, but the helper is beginning to steer. "Why do you suppose that happened?"
- **Confrontive Inquiry**: Embedding your own idea in the form of a question. "Confrontive inquiry is in pursuit of information related to something that you, as the helper, want or that you are thinking about. You have crossed over the net, and it becomes about you as much or more than the client" (§0.0.444).
- **Process-Oriented Inquiry**: Asking about the conversation itself. "What do you think is happening between us right now?" (§0.0.480).

The failure mode that matters most for this review is confrontive inquiry — questions that sound like genuine curiosity but embed a conclusion the asker has already reached. "What is most important is to first confront yourself with the question of what your purpose is before you ask a confrontive question. Are you feeling curious, or have you fallen into thinking you have an answer

and are just testing out whether or not you are right? If you are just testing your own ideas, then you have drifted into telling and it should not be surprising if the other person gets defensive" (§0.0.449).

Schein's most practically important claim: performed curiosity is detectable. "Faux humility comes across loud and clear. Generally, no matter how you phrase your questions, others will sense it immediately if you are not at all interested in them. At the same time, if the attitude behind the behavior is correctly perceived to be sincere interest, even some kinds of telling may have the same positive effect as Humble Inquiry" (§0.0.352). The attitude can carry an imperfect question; the right words cannot carry an insincere attitude.

This is why Schein distinguishes Here-and-now Humility from humility as a personality trait. You do not need to be a humble person to practice humble inquiry. You need to accurately perceive, in the present moment, that you are genuinely dependent on what the other person knows. "Even the most confident or arrogant among us will find ourselves humbled by the reality of being dependent on others... We can think of this as Here-and-now Humility, accepting our dependence on each other for information sharing and task completion" (§0.0.121). When the dependence is real and accurately perceived, the questions that follow it will carry the genuine quality that makes them land differently.

1.7.3 *Conversational Receptiveness (Minson & Yeomans)*

The third posture provides the empirical grounding for why the first two matter. Julia Minson and Michael Yeomans (2021) identified the verbal markers of a disposition to engage genuinely across disagreement — a construct they call **conversational receptiveness** — and tested what happens when it is present or absent.

Receptive responses include: acknowledging the other's perspective ("I can see why you'd think that"), using hedging language that signals appropriate uncertainty ("I may be wrong, but..."), and noting areas of genuine agreement before areas of disagreement. These are not rhetorical flourishes. They are behavioral signals that disagreement will not be met with attack — that it is safe to reveal uncertainty, to acknowledge partial agreement, to say "I'm not entirely sure."

Minson and Yeomans' finding (Organizational Behavior and Human Decision Processes, 2021): people exposed to receptive responses are more willing to continue the conversation, more likely to share their actual views, and more likely to update. The mechanism is straightforward: receptiveness signals psychological safety in the epistemic domain. When a person knows that their uncertainty will not be exploited and their partial concessions will not be treated as total defeats, they can surface the doubts they actually have — which are then available for the questioning moves of the playbook.

This finding provides the empirical grounding for the review's first integrative finding: non-judgmental listening is the universal precondition. Eight source traditions — Motivational Interviewing, Street Epistemology, Harvard negotiation, deep canvassing, Verbal Judo, Humble Inquiry, pragma-dialectics, and Minson's own research — converge independently on this requirement. The convergence is remarkable precisely because these traditions have different origins, different goals, and different methods. They all discovered, through different routes, that you cannot move someone's thinking from the outside. You can only help them move it from the inside — and that requires creating the conditions under which they are willing to actually look at it.

Receptiveness is not the same as agreement. The practitioner who demonstrates conversational receptiveness is not capitulating or performing false open-mindedness. They are signaling that this

is a conversation, not a combat — and that signal makes possible the kind of genuine examination that can sometimes, under the right conditions, move from discussion to dialogue in Bohm's sense.

1.8 THE INTEGRATED MAP: THREE LAYERS, ONE DIAGNOSTIC

Walton, Bohm, and van Eemeren are sometimes presented as competing frameworks for understanding argument and dialogue. They are better understood as operating at different levels of abstraction, with complementary functions.

Use Walton's typology for structural diagnosis: What type of dialogue is this? What are the parties' individual goals, and are they compatible with the conversation's collective goal? Has there been a dialectical shift — and if so, was it an embedding or an illicit derailment?

Use Bohm's quality dimension for phenomenological diagnosis: Is meaning actually flowing? Are the parties genuinely examining the territory, or defending their maps? Has the conversation shifted from discussion mode (win-lose, positions defended) to something closer to dialogue mode (win-win, errors shared as gains)? Is suspension occurring, or is there only suppression?

Use van Eemeren's ten rules for normative diagnosis: Is this participant — am I — behaving in good faith? Which rules are being violated, and what does each violation reveal about what game is actually being played?

The three layers provide the practitioner with a rich diagnostic vocabulary. The composite diagnostic provides the practical triage. The three dispositional postures — scout mindset, humble inquiry, conversational receptiveness — provide the internal conditions under which the diagnostic can be applied honestly rather than deployed as sophisticated ammunition.

The map that results is not a guarantee of productive conversation. It is the equipment for understanding what is happening well enough to make informed choices about what to do next. Those choices — the named moves of the playbook — are the subject of Part II.

Part I word count: approximately 6,200 words.

2

PART II: THE PLAYBOOK

2.1 HOW THIS SECTION WORKS

The playbook that follows is organized by function, not by source. The same move — mirroring, for instance — appears in hostage negotiation under Voss's name, in Motivational Interviewing as part of the OARS framework, in Verbal Judo as a step in the LEAPS sequence, and in Street Epistemology as a technique for keeping the other person talking. These are not four different moves. They are the same move discovered independently across four traditions because it solves the same problem at the same moment in a conversation. The organizing principle here is not authorship but function: what are you trying to accomplish right now, and which move does that job?

Six functional categories correspond roughly to the phases of a real conversation. Listening moves establish the precondition — they create the psychological safety without which no other technique operates. Questioning moves surface the other person's epistemology without confrontation — they ask how a belief was formed, not whether it is correct. Reframing moves shift the conversation's frame: from positions to interests, from blame to contribution, from debate to shared problem. Face-saving moves build the exit ramp — they allow position change without humiliation, which cognitive dissonance research shows is the primary barrier to belief revision. Diagnostic moves help you recognize, in real time, what game you are in and whether joint inquiry is achievable at all. Disengagement moves describe when and how to stop.

The sequencing is loose but not arbitrary. In a difficult conversation you typically need to listen before you can question, question before you can reframe, reframe before you can build a face-saving exit. But conversations are not linear, and the playbook should be read as a reference, not a script. Someone in the middle of a real exchange should be able to scan by heading, find the relevant move, and have the script visible without reading the surrounding commentary.

Three integrative findings from the research literature run through every section and justify the architecture. First: non-judgmental listening is the universal precondition across eight independent traditions. Second: narrative exchange outperforms argument for shifting entrenched views — sharing personal experience and inviting the other person's story is more effective than presenting evidence (Broockman & Kalla, 2016; Kalla & Broockman, 2020). Third: technique rebuttal outperforms topic rebuttal — examining how a belief is held is more effective than arguing about what the belief says (Schmid & Betsch, 2019, PNAS). Every move in the playbook is downstream of one or more of these findings.

3

1. LISTENING MOVES

Non-judgmental listening is not a soft nicety. It is the empirical precondition for everything else in this playbook. Eight traditions arrived at this finding independently: Motivational Interviewing, Street Epistemology, Verbal Judo, the Harvard Negotiation Project, deep canvassing, pragma-dialectics, Bohm's dialogue theory, and Schein's humble inquiry. The convergence is striking enough that the finding can be stated with unusual confidence: no other technique operates without it.

What the research specifies is more precise than "listen carefully." Santoro, Broockman, Kalla, and Porat (2025, PNAS) found that high-quality listening improves the other person's perception of the listener but does not itself change attitudes. Listening is not the intervention — it is what makes the intervention possible. The persuasive effect, when it occurs, comes from the narrative that the listening unlocks: the person, feeling safe, reveals something genuine. Itzchakov and DeMarree (2022) demonstrated that perceived listening quality predicts attitude change via psychological safety — the interlocutor becomes willing to reveal their doubts only when they are not afraid of being attacked for having them. Doubt, not certainty, is the material the rest of the playbook needs to work with. Listening is what surfaces it.

3.1 MOVE 1.1 — REFLECTIVE LISTENING

Source: Motivational Interviewing (Miller & Rollnick, 2013, OARS framework)
Native type: All types. Universal entry move.

> **"It sounds like you're saying..." / "What I'm hearing is..." / "So on one hand [position], and at the same time [doubt or concern]..."**

Rather than responding to the content of what someone said, reflect back the meaning — both what was stated and what was implied. Miller and Rollnick distinguish three levels, each progressively more powerful. A **simple reflection** repeats or lightly rephrases what was said, demonstrating basic comprehension. A **complex reflection** paraphrases with emotional content added — it names what seems to be behind the words, not just the words themselves. A **double-sided reflection** is the most sophisticated: it names both the stated position and the ambivalence behind it simultaneously: "On one hand, you believe X, and at the same time, you seem uncertain about Y." The double-sided reflection is the form closest to Bohm's dialogue ideal — it holds the tension rather than collapsing it.

Reflective listening fails when it is done mechanically. Mere parroting — repeating back words without genuine understanding — feels like mockery. The move works only when the listener is actually trying to understand, and the person can tell the difference. This is why Galef's scout mindset (Galef, 2021) and Schein's humble inquiry (Schein, 2013) are not optional complements to listening technique — they are its precondition. The words are the same; the orientation is not.

Why it works: Wampold's (2015) meta-analysis of psychotherapy outcome research found that the therapeutic alliance — built substantially on this kind of engaged listening — predicts outcomes better than any specific technique, with an effect size of approximately d = 0.57. Aristotle identified the same mechanism two millennia earlier: ethos precedes logos. You cannot persuade someone

who does not trust you. Reflective listening is how trust is established quickly in an unfamiliar conversation.

Failure mode: Mechanical repetition without genuine curiosity creates a robotic, interrogatory atmosphere. Failure reveals that the listener does not actually want to understand — they are performing interest. This is detected immediately and is often more damaging than open disagreement.

Hard case applicability: All three — anti-intellectualism, high-conflict personalities, narrative-pushers. Reflective listening is the baseline entry move regardless of context.

3.2 MOVE 1.2 — MIRRORING

Source: Hostage negotiation (Voss, 2016); also Boghossian & Lindsay (2019)
Native type: Negotiation, persuasion → shift target: information-seeking, inquiry

> **Repeat the last 2–3 words of what the person said as a question.** *(pause — let silence do the work)*

The mechanism is simple and counterintuitively powerful. People feel compelled to elaborate when their own words are returned to them as an open question. Voss, drawing on FBI hostage negotiation practice, reports that a well-timed mirror extends the other person's disclosure reliably. The pause after the mirror is load-bearing — breaking it too quickly forfeits the effect.

Voss suggests approximately four to five mirrors per seven-minute conversation. Used more often, the technique becomes obvious and loses its quality of genuine curiosity. Paired with silence, it generates information without commitment: you learn what the person actually thinks rather than what they want you to know they think.

Failure mode: Overuse produces a robotically interrogatory atmosphere. The move also fails when the person shifts topics quickly — mirroring a dead branch of conversation generates confusion rather than elaboration. Failure reveals that the person has run out of things to say on this branch, or has become suspicious that they are being "managed."

Hard case applicability: Most useful with narrative-pushers (they will elaborate the point they are trying to make) and moderately useful with anti-intellectualists (they will extend their account of why expertise cannot be trusted). Caution with high-conflict personalities — mirroring can extend rather than redirect hostile monologue.

3.3 MOVE 1.3 — LABELING

Source: Hostage negotiation (Voss, 2016); Verbal Judo (Thompson & Jenkins, 2013)
Native type: Negotiation, eristic-approaching → shift target: persuasion

> **"It seems like you feel pretty strongly that this is unfair." "It sounds like this has been really frustrating."** *(pause for correction)*

Name the emotion you observe beneath what is being said. The phrasing "it seems like" or "it sounds like" is load-bearing: it offers the move as a hypothesis, not a declaration. The person can correct your label, which is itself useful — the correction reveals what they actually feel. Either way, the act of naming the emotion demonstrates that you noticed it, and noticing it is most of what the person needs before they can engage with anything else.

Kahneman's System 1/System 2 distinction (2011) provides the explanatory mechanism. Emotional pressure must be addressed before analytical processing can operate reliably. Presenting a logical argument to someone who feels unheard is addressing System 2 while System 1 is running interference. The label addresses System 1 directly — it says, in effect, "I see that you are experiencing something, and I am not pretending otherwise."

Failure mode: Naming the wrong emotion, or naming the right emotion too explicitly when it has not yet been fully acknowledged, triggers denial and defensiveness. "You seem angry" when someone has not yet admitted they are angry forces them to either confirm (exposing themselves) or deny (closing the channel). Softer hypothesis framing avoids this. Failure reveals that the conversation needs more reflective listening before labeling.

Hard case applicability: Essential in high-conflict personality encounters, where the emotional content is often the entire substance of the exchange. Useful in anti-intellectualist encounters where the underlying feeling is often one of being condescended to by expertise. Less central in narrative- pushing encounters, where the person is typically not in an emotionally activated state.

3.4 MOVE 1.4 — LEAPS SEQUENCE

Source: Thompson & Jenkins, Verbal Judo (2013)

Native type: Eristic-approaching, high-conflict → shift target: persuasion, information-seeking

"Let me make sure I understand. You're saying that [paraphrase]. Is that right?"

The LEAPS sequence is a structured five-step listening protocol designed for high-conflict encounters where the other person has been speaking at you rather than with you:

L — Listen without interrupting, even when you disagree with what is being said. Interruption closes the loop; silence keeps it open.

E — Empathize with the feeling, not necessarily the position. "I can understand why you'd be upset about that" does not endorse the content; it acknowledges the experience.

A — Ask an open question to learn more. The question signals that the listening was genuine, not performative, because genuinely confused people ask questions.

P — Paraphrase to demonstrate understanding. This is where the paraphrase from Move 1.1 is deployed — showing, not merely claiming, that you followed.

S — Summarize what matters. Condensation is an act of interpretation: choosing what to carry forward signals what you understood to be the point.

The sequence slows the pace of a high-conflict encounter, which is itself often a de-escalation. People who feel that they must speak loudly and quickly to be heard slow down when the listener is not rushing to respond. Thompson and Jenkins describe this as "absorbing the force" of the other person's communication — yielding without retreating.

Failure mode: Any step that is rushed or perfunctory destroys the sequence's value. A quick "I hear you" followed immediately by a pivot to your own point does not constitute empathy and will be recognized as an evasion. Failure reveals that you are performing patience rather than exercising it.

Hard case applicability: Designed for high-conflict personalities and high-emotionality encounters. Less necessary in calm persuasion dialogue.

3.5 MOVE 1.5 — CONVERSATIONAL RECEPTIVENESS SIGNALING

Source: Minson & Yeomans (2021, OBHDP); Minson, Hagmann & Luo (2023)

Native type: All types, opening

> **"I can see why that would be your view, given [reason]. I'm less sure about [specific point] — can you help me understand [specific question]?"**

Minson's research identifies specific linguistic markers that signal genuine engagement to the other party, increasing their willingness to continue the conversation and reveal their actual views rather than a performed position. The markers are: acknowledging their perspective ("I can see why you'd think that"), hedging your own claims ("I may be wrong, but..."), explicitly noting genuine points of agreement before any disagreement, and expressing uncertainty as invitation rather than weakness ("I'm not sure I follow — can you say more about that?").

The mechanism is psychological safety, and it is specific: if disagreement will be met with attack, people perform certainty. When it is safe to reveal doubt, they reveal it. And doubt — the other person's own uncertainty about their own position — is the material the questioning moves in the next section need to work with. Receptiveness signaling creates the conditions for uncertainty to surface.

Failure mode: Performed receptiveness — saying the words without the underlying orientation — is detected rapidly and discredits everything that follows. The person who has just been told "that's a really interesting perspective" when you clearly do not find it interesting experiences the condescension of insincere affirmation. Failure reveals that your curiosity is tactical rather than genuine, which forecloses the conversation at a deeper level than simple disagreement would have.

Hard case applicability: Particularly valuable in anti-intellectualist encounters, where the person has often been condescended to by people who hold evidence-based positions and has developed a sensitive detector for it. Genuine receptiveness is precisely the opposite of what they expect.

4

2. QUESTIONING MOVES

The central integrative finding of the research literature — that technique rebuttal outperforms topic rebuttal — is operationalized in this section. Schmid and Betsch (2019, PNAS) found that interventions that examine how a belief is held are more effective than those that argue against what the belief says. Every move in this section asks some version of HOW: How did you come to think that? How confident are you, and why not more? How would you know if you were wrong? Arguing against WHAT someone believes forces them to defend it — the righting reflex, in Motivational Interviewing's vocabulary. Asking about HOW they came to believe it invites them to examine their own epistemic process from the inside, in a context where they are the expert on their own reasoning.

Mercier and Sperber's argumentative theory of reasoning (2017) provides the explanatory backdrop. If reasoning evolved as a tool for winning arguments rather than for tracking truth, then the standard adversarial approach — present evidence, expect the other person to update — is naïve about the cognitive apparatus it is engaging. The questioning moves work with that apparatus rather than against it.

4.1 MOVE 2.1 — THE EPISTEMOLOGICAL SEQUENCE (STREET EPISTEMOLOGY)

Source: Boghossian & Lindsay (2019); Magnabosco SE corpus
Native type: Persuasion → shift target: inquiry

> **Step 1:** "On a scale of 1 to 10, how confident are you in that?" **Step 2:** "Why not a [higher number than they gave]?" **Step 3:** "How did you come to be confident in that?" **Step 4:** "Under what conditions could that belief be wrong?"

The sequence is designed to move from confidence to epistemology without triggering the defensive response that direct challenge provokes. Step 1 establishes that confidence is a variable rather than a binary — it creates conceptual space for the person to locate themselves somewhere other than "completely certain." Most people, when asked to name a number, choose something below 10, which means they have already implicitly acknowledged doubt.

Step 2 is the key move. Asking why they did not choose a *higher* number is counterintuitive: it redirects attention toward their existing doubts rather than asking them to generate new reasons for their position. The doubts are already there — this question makes them explicit and conversationally available.

Step 3 is the epistemological question proper. It shifts from the content of the belief to its formation: not "why do you believe X?" (which typically elicits rehearsed arguments in favor of X) but "how did this belief come to seem credible to you?" The latter question invites reflection on the process of belief formation — on what counted as evidence, whose testimony was trusted, what was considered and what was not.

Step 4 is the disconfirmation question, treated in detail in Move 2.2 below.

Failure mode: If any step in the sequence is rushed, or if the listener responds to the person's answer with their own view rather than another question, the sequence collapses into a standard

argument. The questioner must remain genuinely curious — Galef's scout posture — throughout. Failure reveals that the questioner was using the form of inquiry while actually doing persuasion.

Empirical status note: Street Epistemology as a packaged technique has no peer-reviewed randomized controlled trial evaluation. The supporting evidence is indirect: the non-confrontational questioning approach is supported by MI literature (Lundahl et al. 2010); the focus on epistemological mechanisms over content is formalized in Schmid & Betsch (2019); narrative elicitation has RCT support via Broockman & Kalla (2016). The SE community has produced over 1,000 documented conversations (Magnabosco's corpus) but without control conditions or outcome measurement.

Hard case applicability: Highly applicable to anti-intellectualism (epistemology of why expertise is distrusted) and narrative-pushing (the person's confidence rating for their own point is typically unstable on examination). Less applicable to high-conflict personalities, where the sequence will often be derailed at Step 1.

4.2 MOVE 2.2 — THE DISCONFIRMATION QUESTION

Source: Boghossian & Lindsay (2019, §0.60.1)
Native type: Persuasion → shift target: inquiry

> **"Under what conditions could that belief be wrong?" "What evidence, if you encountered it, would cause you to update?"**

This is the most diagnostically powerful question in the playbook. The answer reveals not merely the content of the belief but its epistemic structure — specifically, whether it is held as a claim about the world (disconfirmable in principle) or as a commitment (held independently of evidence). Three response categories, each with distinct implications:

Genuinely disconfirmable: The person names specific conditions — "I'd update if I saw X." Explore what X looks like and whether it exists — but do not deliver it yourself (see Move 2.5 on outsourcing). The goal is to elicit the person's own search, not to introduce evidence that will be treated with suspicion because of its source.

Implausibly disconfirmable: "I'd update if aliens turned out to be controlling the media" — a nominally disconfirmable condition that functions as non-disconfirmable. Test the sincerity of the stated conditions: "What would you need to see to believe there are no aliens controlling the media?" The conditions become recursive or absurd.

Non-disconfirmable: "Nothing could make me wrong about this." This is the most diagnostic response of all. It reveals that the belief is not an empirical claim but a commitment held on identity or moral grounds. Pivot to Move 2.3.

One critical note: Boghossian and Lindsay (2019) invoke the "backfire effect" — the claim that correcting a false belief causes the person to hold it more strongly — to justify the indirect approach of eliciting doubts rather than presenting evidence. This claim has not replicated robustly. Wood and Porter (2019) failed to find the backfire effect across 52 belief-correction experiments covering 16 political issues. Ecker, Lewandowsky, and colleagues (2022, Nature Reviews Psychology) is the current authoritative synthesis: the backfire effect is not a reliable phenomenon. What IS established is that corrections are often insufficient for beliefs held on identity grounds — they do not work as reliably as one might hope. What is NOT established is that corrections reliably make false beliefs stronger. The practical advice — elicit the person's own doubts rather than delivering evidence directly — survives this revision, but for a more nuanced reason than Boghossian and Lindsay state. The elicitation approach works not because evidence causes backfire but because

self-generated doubt is more persuasive than externally delivered evidence, consistent with the Kurt Lewin findings that anchor Motivational Interviewing.

Failure mode: Asking the disconfirmation question before sufficient rapport and epistemic trust have been established triggers defensiveness. The question can sound like a trap. Failure reveals that the listening moves (Section 1) have not yet done their work.

Hard case applicability: Most useful with anti-intellectualism and narrative-pushing. The diagnostic value alone justifies its use — if the answer is non-disconfirmable, you have learned something important about how to proceed.

4.3 MOVE 2.3 — THE IDENTITY QUESTION

Source: Boghossian & Lindsay (2019, §0.62.5–§0.62.8)
Native type: Persuasion → moral/identity territory

> **"What kind of person holds that belief?" "What would it mean about you if you changed your mind?" "Can you think of good people who don't hold this view?"**

When the disconfirmation question has established that a belief is held on moral or identity grounds — when "nothing could make me wrong" is the honest answer — the belief is functioning as what Daniel Dennett called a "belief in a belief": the person believes they believe X because believing X is what makes them a good person, a loyal member of their community, or an epistemically serious person. Arguments at the level of the belief will fail because the defenses operate at the level of identity.

The identity questions shift the conversation to that level. They are not rhetorical jabs — they are genuine inquiries into the relationship between the belief and the person's self-conception. Most people, when sincerely asked whether good people can disagree with their position, will pause. That pause is cognitively significant: it represents actual consideration of whether the belief is constitutive of goodness or merely associated with it.

Boghossian and Lindsay (2019) offer an "Ideologue Template" for sustained cases where the identity questions are also resisted: "It's clear to me that being a good person is important to you. What values underlie that belief? How do you know that acting in that way makes someone good?" This is a move to the level of moral epistemology — how does the person know what counts as good? Defenses at the belief level are well-rehearsed; defenses at this deeper level are typically more fragile.

Failure mode: The identity question can feel invasive or condescending if the person does not believe their identity is implicated. Verify that the belief is genuinely being held on identity grounds (via Move 2.2) before deploying this move. Failure reveals that the question was deployed too early or at the wrong level.

Hard case applicability: Central to anti-intellectualism encounters (where distrust of expertise is often a deeply identity-relevant commitment), and possibly useful with narrative-pushers. Not recommended for high-conflict personalities, where questions about identity will often be redirected into personal attack.

4.4 MOVE 2.4 — OPEN QUESTIONS (MOTIVATIONAL INTERVIEWING OARS)

Source: Miller & Rollnick, Motivational Interviewing (2013)
Native type: Information-seeking → persuasion

> **"What concerns you most about [topic]?" "What would it mean for you if [change]?" "What do you already know about [topic]?"**

The OARS framework — Open questions, Affirmations, Reflective listening, Summaries — is the operational core of Motivational Interviewing. The open question is the inquiry instrument. Miller and Rollnick draw a sharp distinction between closed questions (which elicit yes/no or single-item answers and transfer control to the questioner) and open questions (which invite elaboration and leave control with the person).

The central MI insight — what they call the "righting reflex" — explains why the direction of inquiry matters. The righting reflex is the natural impulse to correct someone who seems to be going wrong: to tell them the right answer, the correct information, the better path. When the practitioner expresses the case for change, the person argues against it. When the practitioner elicits the case for change from the person, the person argues for it. This is why the open question — "what concerns you about this?" rather than "here's why you should be concerned about this" — is so consistently more effective than the lecture.

The empirical foundation for MI is the most robust of any technique in this review. Miller's original paper dates to 1983; Lundahl, Kunz, Brownell, Tollefson, and Burke's (2010) meta-analysis covering twenty-five years of randomized controlled trials found effect sizes of g = 0.28 compared to no treatment, across substance abuse, health behavior, and treatment adherence. Smedslund and colleagues' (2011) Cochrane review confirms. MI was developed in clinical substance abuse treatment, but the open-question mechanism is general.

Failure mode: The righting reflex is hardest to suppress precisely when the other person is most wrong. The practitioner who is genuinely trying to ask open questions will feel the pull to deliver the answer rather than elicit it. Failure reveals the practitioner's own righting reflex — a diagnostic about the listener, not the speaker.

Provenance note: *Motivational Interviewing* (3rd ed., 2013) was not available in the Reading Guides pipeline. The OARS framework, righting reflex concept, and scripts above are reconstructed from Miller & Rollnick (2009) "Ten Things MI Is Not" (Abstract-verified), Lundahl et al. (2010) meta-analysis (Abstract-verified), and training knowledge. Verbatim scripts should be verified against the primary text if the PI obtains it.

Hard case applicability: All three, but the mechanism is most clearly applicable to anti-intellectualism (eliciting the person's own concerns about the positions they hold) and narrative-pushing (eliciting what the person is actually trying to say).

4.5 MOVE 2.5 — OUTSOURCING (EVIDENCE ELICITATION)

Source: Boghossian & Lindsay (2019, §0.51.2–§0.51.5)
Native type: Persuasion → inquiry

> **"Where would you look to find out more about this?" "Who are the three best experts who disagree with this position?" "Who's the strongest voice on the opposing side?"**

When a conversation has produced a specific empirical question — one where evidence could in principle settle the matter — the impulse is to deliver the evidence. Resist this impulse. Outsourcing asks the person to identify where they would look for information, rather than receiving information from you.

Two mechanisms explain why this works. First, the messenger effect: people resist information that arrives from sources they perceive as motivated to influence them. Evidence delivered by an

opponent in an argument is maximally suspicious regardless of its quality. Evidence the person locates themselves, via sources they already find credible, arrives without this suspicion. Second, the epistemic trust question is surfaced: "Who are the best experts who disagree with you?" asks the person to engage with their own epistemic trust structure — which sources count as credible to them, and why. This is preparation for later work at the epistemological level.

This move is only applicable to empirical claims. Do not outsource on moral questions — "go find the best argument for the opposing moral position" is structurally different from "go find the best evidence against your empirical claim," and conflating them will be experienced as a trap.

Failure mode: Asking too early, before the person has any felt uncertainty about their position, produces dismissal. "I know where to look, and I already have" is the response when doubt has not yet been opened. Failure reveals that the disconfirmation questions (Move 2.2) have not yet done their work.

Hard case applicability: Anti-intellectualism encounters, where the question "who would you trust to disagree with you?" surfaces the epistemic trust structure directly. Less applicable to high-conflict personalities.

4.6 MOVE 2.6 — HUMBLE INQUIRY (DEFAULT POSTURE)

Source: Schein, Humble Inquiry (2013)

Native type: All types, as default opening posture

> **"What's your sense of [topic]?" "Help me understand how you think about this." "What am I missing here?"**

Humble inquiry, in Schein's formulation, is the practice of asking questions to which you genuinely do not know the answer, while suppressing the impulse to embed your agenda in the question form. "Why did you do that?" is not humble inquiry — it contains the implication that there may be no good reason. "What were you trying to accomplish?" is humble inquiry because it does not contain that implication.

Schein distinguishes humble inquiry from diagnostic inquiry ("How do you feel about that?" — still agenda-neutral but directional toward feelings) and from confrontational inquiry ("Didn't you realize that was a mistake?" — which contains the agenda in the question). Humble inquiry is prior to all of these: it is the foundational posture of genuine curiosity rather than information extraction.

The connection to Galef's scout mindset (2021) is direct. The scout asks questions because they want to know what the territory looks like. The soldier asks questions to find the opponent's weaknesses. The verbal forms of the questions can be identical — "why do you believe that?" — while the underlying motivation is opposite. The difference in motivation is perceptible to the person being questioned, and it shapes their response.

Failure mode: Humble inquiry is less a move than a posture, and postures cannot be performed without being adopted. The failure mode is not specific phrasing — it is the orientation behind the phrasing. Asking humble-sounding questions while internally planning the rebuttal is not humble inquiry. Failure reveals that the practitioner has not adopted the scout mindset and cannot make this move available to themselves.

Hard case applicability: Anti-intellectualism encounters, where genuine curiosity about why someone distrusts expertise is both rare (the person has rarely encountered it) and effective (it violates their expectations in a productive way). Narrative-pushing encounters, where humble inquiry is the primary tool for discovering what the person is actually trying to say.

5

3. REFRAMING MOVES

Reframing is not spin. It is not finding a less offensive way to say the same thing. It is the move from one frame — positions, blame, debate — to a different frame — interests, contribution, shared problem — that makes previously impossible movement possible. When two people are arguing from entrenched positions, no amount of argument within those positions will produce movement. The positions are the problem. Getting underneath them, or beside them, or to a higher level of abstraction where they can be examined rather than inhabited, is the work of reframing.

The Harvard Negotiation Project literature — Fisher and Ury (2011), Ury (1991), Stone, Patton, and Heen (2010) — provides the primary techniques here. But the underlying logic runs through Bohm's instruction to "suspend assumptions," van Eemeren's normative framework for good-faith dialogue, and Perelman's concept of dissociation. These are not merely negotiation techniques applied beyond their domain — they are applications of a general principle: most conversational impasses are not about the stated subject but about a frame that has made the stated subject unresolvable.

5.1 MOVE 3.1 — INTERESTS BEHIND POSITIONS

Source: Fisher & Ury, Getting to Yes (2011)
Native type: Negotiation → deliberation → persuasion

> **"It sounds like what you're really concerned about is [interest]. Is that right?" "Why is [position] important to you?" "Help me understand what would need to be true for you to feel satisfied here."**

Any stated position — "I want X" — conceals an underlying interest: "because I need Y." The position is one possible solution to the interest; it is rarely the only one. By moving from positions to interests, you create solution space that positional bargaining would never find.

Fisher and Ury's example remains canonical. Two parties argue over who gets a single orange. If they divide it down the middle, one person gets half to eat and half a rind they don't need; the other gets half a rind for a recipe and half an inside they don't need. Stated positions (each wants the whole orange) concealed non-competing interests (eating the flesh vs. using the rind) that a simple interests question would have revealed.

The move is not restricted to explicit negotiations. In any argument where two people hold incompatible positions, the interests question asks what each position is in service of — and interests are often either compatible or resolvable at a level of abstraction where the positions are not.

Fisher and Ury's BATNA concept is worth naming here: your Best Alternative to a Negotiated Agreement is the walk-away threshold that should be established before any negotiation. Without knowing your BATNA, every offer looks acceptable relative to nothing. With it, you can evaluate proposals against a genuine alternative rather than in isolation.

Failure mode: Misnaming the interest — assuming you know why someone wants what they want rather than asking — triggers defensive correction. Failure reveals that insufficient listening preceded this move.

Hard case applicability: Narrative-pushing encounters, where the move clarifies what the person is actually trying to argue beneath the argument they are making. Moderate applicability to anti-intellectualism, where the stated position (distrust of experts) often conceals an interest in being treated as an autonomous epistemic agent.

5.2 MOVE 3.2 — GO TO THE BALCONY

Source: Ury, Getting Past No (1991)
Native type: All types, as metacognitive rescue move

> **"That's an interesting point. Let me think about that for a moment." "I'd like to take a pause here — I want to make sure I'm understanding you correctly."**

When you feel triggered, provoked, or pulled into eristic mode, mentally step to the balcony — Ury's image for the deliberate pause to observe the situation from above rather than react from within it. From the balcony, you can see the conversation as a system rather than experiencing it as a participant. From inside it, every provocation demands a response; from the balcony, you can choose whether to respond and how.

The balcony move is a unification point across traditions. Walton's concept of "suspension" — the willingness to hold your assumptions rather than act from them — is the same move from argumentation theory. Bohm's instruction to suspend your thoughts "in front of you" as objects to be examined rather than positions to be defended is the same move from dialogue philosophy. Ury's balcony is the same move in negotiation practice. Three traditions, one cognitive operation: metacognitive observation of your own reasoning in real time.

In practice, the balcony is accessed via any move that buys time without surrendering momentum. The scripts above serve this function — they are genuine expressions of the pause rather than stalling tactics, which is why they must be sincere.

Failure mode: The move fails when it becomes avoidance — taking refuge in "let me think about that" as a way of never responding. The balcony is a pause, not an exit. It fails also when the practitioner descends from the balcony with the response they would have given without taking the pause — meaning they used the pause to recover emotionally without using it to reconsider.

Hard case applicability: Essential for high-conflict personality encounters and any encounter where the other person is using provocation strategically. The refusal to be provoked is itself a move — it denies the other party the satisfaction of the expected reaction.

5.3 MOVE 3.3 — THE FIVE-STEP BREAKTHROUGH STRATEGY

Source: Ury, Getting Past No (1991)
Native type: Eristic-approaching → negotiation → persuasion
A structured sequence for encounters that have become entrenched or hostile:

> **Step 1 — Go to the Balcony:** Control your own reaction. (Move 3.2) **Step 2 — Step to Their Side:** "I can see why you'd see it that way." (Disarm without agreeing) **Step 3 — Reframe:** "What are we actually trying to accomplish here?" (Move from positions to interests) **Step 4 — Build a Golden Bridge:** "How could we make this work for both of us?" (Make agreement easy — see Move 4.1) **Step 5 — Make It Hard to Say No:** "What happens if we can't reach agreement?" (Logical consequences, not threats)

Step 2 is the most technically demanding. "Stepping to their side" means acknowledging the other person's perspective, understanding their case, and recognizing their feelings — without endorsing their position. The phrasing "I can see why you'd see it that way" is precisely calibrated: it is genuine acknowledgment of their epistemic situation without agreement with their conclusion.

This step is also the point where Ury's framework converges with multiple other traditions. It is Rapoport's first rule from argumentation theory: before criticizing a position, state it clearly enough that its proponent would recognize it. It is Motivational Interviewing's reflective listening. It is Boghossian and Lindsay's "express empathy" opening from their Socratic dialogue protocol. Three traditions, the same move, for the same reason: a person who does not feel heard cannot update.

Failure mode: Step 2 fails when "I understand your perspective" is used as a setup for "but here's why you're wrong" — which the other person will detect immediately and treat as a particularly dishonest form of dismissal. The sequence requires genuine engagement at each step.

Hard case applicability: Designed specifically for eristic-approaching and high-hostility encounters. Applicable to all three hard cases.

5.4 MOVE 3.4 — THE CONTRIBUTION FRAME

Source: Stone, Patton & Heen, Difficult Conversations (2010)
Native type: Persuasion, negotiation → avoids eristic

> **"What factors contributed to [situation]?" "How might each of us have contributed to where we are?" "What could I have done differently?"** *(acknowledge your own contribution first)*

Stone, Patton, and Heen identify three simultaneous conversations in any difficult exchange: the "What happened?" conversation (facts and interpretations), the "Feelings" conversation (emotions that haven't been explicitly named), and the "Identity" conversation (what this situation means about who I am). Their contribution frame applies specifically to the first conversation, where the instinct is to assign blame.

Blame is binary, backward-looking, and triggers defensiveness because it assigns moral responsibility and therefore threatens identity. The contribution question replaces blame with a systemic analysis: what factors, including the behavior of both parties, produced this outcome? Contribution is pluralistic — many factors contribute, and acknowledging one's own contribution does not mean accepting total responsibility. It is also forward-looking in implication: understanding contribution patterns makes it possible to change them.

Acknowledging your own contribution first, before asking about the other person's, is critical. It demonstrates that the frame is not a rhetorical device for extracting an admission — it is a genuine shift in how the situation is being analyzed.

Failure mode: If the contribution frame is introduced before rapport exists, it can sound like blame in disguise: "How did you contribute to this?" reads as an accusation regardless of framing when the relationship is already adversarial. Failure reveals that the listening moves have not preceded this move.

Hard case applicability: Most useful in narrative-pushing encounters where the person is pushing a narrative of pure victimhood or pure responsibility. Moderate applicability to anti-intellectualism. Caution with high-conflict personalities, for whom contribution framing may extend rather than resolve the conflict.

5.5 MOVE 3.5 — DISSOCIATION

Source: Perelman & Olbrechts-Tyteca, The New Rhetoric (1969)
Native type: Persuasion → inquiry

> **"I wonder if we're conflating two things here — [concept A] and [concept B]. Could we look at those separately?" "Is this [X] in the sense of [narrow definition], or [X] in the sense of [broader definition]?"**

Perelman and Olbrechts-Tyteca identified dissociation as one of the most powerful moves in argumentation: the separation of concepts that have been fused in ways that foreclose productive discussion. The classic form is the appearance/reality distinction: "that may be what appears to be justice, but is it real justice?" Dissociation acknowledges that the term is being used while questioning which use is operative.

In practice, dissociation is most useful when a conversation is stuck because a loaded term is doing too much work. "Freedom" can mean the absence of external constraint, the presence of real options, or a political identity. "Fairness" can mean equal treatment, proportional treatment, or outcomes. When two people use the same word with different meanings and neither notices, the argument proceeds as if they disagree about the world when they may disagree only about the word. Dissociation makes the distinction explicit and allows each sense to be evaluated independently.

Failure mode: Introduced too aggressively, dissociation sounds like "you don't know what you mean" — which is condescending. The move requires a light touch: "I wonder if..." rather than "You're confusing...". Failure reveals insufficient attention to how the move will be received.

Hard case applicability: Particularly valuable in anti-intellectualism encounters, where terms like "expert," "evidence," and "consensus" are often used with conflated meanings. Useful in narrative-pushing encounters where the person's point becomes clearer once you help them disentangle the concept they are trying to use.

5.6 MOVE 3.6 — MORAL REFRAMING

Source: Haidt, The Righteous Mind (2012); Boghossian & Lindsay (2019, §0.99.1–§0.100.5)
Native type: Persuasion, cross-value disagreements

> **"I care about this for [reason that maps onto their values], not just [reason that maps only onto mine]."**

Haidt's Moral Foundations Theory identifies six foundational axes on which moral intuitions are organized: care/harm, fairness/reciprocity, loyalty/betrayal, authority/subversion, sanctity/degradation, and liberty/oppression (Haidt, 2012). Research on political psychology suggests that liberal-leaning people tend to weight the first two most heavily; conservative-leaning people tend to weight all six more evenly, with loyalty, authority, and sanctity carrying substantial weight. Cross-value disagreements are often disagreements about which foundations are operative, not about the foundations themselves.

Moral reframing translates an argument into the moral vocabulary of the person you are speaking with — not by abandoning your own values, but by finding the version of your concern that connects to theirs. A concern about healthcare access that is framed entirely in terms of harm prevention (care/harm) may be received very differently when framed also in terms of fairness across contribution levels (fairness/reciprocity) or community obligation (loyalty).

The critical condition: this move requires genuine fluency in the other person's value system. It cannot be a performance. People detect instrumental reframing — the use of their moral vocabulary as a rhetorical tactic while your underlying concern is unchanged — and detect it reliably. The scout mindset is the precondition: you must actually care about understanding why they weight the foundations the way they do, not merely learn to speak their language tactically.

Failure mode: Performed reframing is detected and discredits everything that follows. It also functions as a form of manipulation that is, in Haidt's framework, an appeal to the elephant while bypassing the rider — it may produce agreement without understanding, which is not joint inquiry. Failure reveals that the practitioner has not adopted the scout mindset.

Hard case applicability: Primarily applicable to anti-intellectualism encounters where the resistance to evidence is organized around moral identity rather than epistemology. Moderate applicability to narrative-pushing.

6

4. FACE-SAVING MOVES

Tavris and Aronson's cognitive dissonance research (2020) documents the pyramid of choice: two people who begin in similar positions, each making small self-justifying decisions, end at opposite extremes. The mechanism is self-justification — each decision requires justifying itself, and the justification shapes the next decision. The result is that changing your mind, by the time the positions have hardened, requires not just updating a belief but unwinding an architecture of self-justification that has been built over time.

This is why the logic-and-evidence approach so often fails on entrenched beliefs: the barrier is not that the person lacks the evidence but that accepting the evidence requires dismantling a scaffolding that has become load-bearing for their sense of themselves as a coherent, good, epistemically serious person. In Storr's framework (*The Status Game*, 2021), being publicly wrong is a status loss — one that most people will do considerable cognitive work to avoid.

Face-saving moves are the mechanism for converting status threats into exits that preserve dignity. They are not dishonest flattery. They are the recognition that position change requires an exit ramp, and that building the ramp is the practitioner's responsibility, not the other person's.

6.1 MOVE 4.1 — THE GOLDEN BRIDGE

Source: Ury, Getting Past No (1991); Boghossian & Lindsay (2019, §0.41.10–§0.41.13)
Native type: All types, as exit-enabling move

> **"Everyone makes mistakes." "This is genuinely complicated — there's a lot of confusion around it." "Expertise is the result of having made many mistakes and changed one's mind accordingly." "Yeah, that concern made total sense given what was publicly known at the time."**

When someone attacks your position personally, recast the attack as about the issue rather than about you, and redirect toward shared values:

> **"I can see why my position might come across that way. I want to be clear that I share your concern about [shared value] — how could we solve that problem together?"**

When someone seems on the edge of updating — when doubt has visibly surfaced — build the exit ramp explicitly. Acknowledge why their previous position was reasonable (it often was, given what they knew or who they trusted). Use "we" and "together" language that converts the conversation from a debate to be won to a problem to be solved. Accept changed minds without reproach, without "I told you so," without triumph of any kind.

The exit must be free of tolls. The moment you extract any form of satisfaction from the other person's update — any visible pleasure at being right — you have destroyed the value of the move and taught them that updating is punished. Ury describes the goal as building a "golden bridge" — wide, solid, and in the direction they can walk without embarrassment.

Tavris and Aronson's pyramid of choice explains why this move is necessary rather than merely kind. By the time a position has hardened, the person is not just updating a belief — they are

dismantling a structure of self- justification that has become load-bearing. The golden bridge does not make that process painless; it makes it possible.

Failure mode: The failure mode is precisely the impulse that makes the move feel unnecessary: the impulse to register that you were right. A toll on the bridge — any form of "I told you so," or "it's about time" — converts the golden bridge into a humiliation device and guarantees that future updates will be more costly. Failure reveals the practitioner's investment in being right rather than in being useful.

Hard case applicability: All three hard cases, but most critically in high-conflict personality encounters, where any perceived "win" by the other person will be met with escalation rather than resolution.

6.2 MOVE 4.2 — BIFF RESPONSE

Source: Eddy, BIFF: Quick Responses to High-Conflict People (2014)
Native type: Eristic, HCP encounters → contains without engaging

> [**One to three sentences. Answer only what was asked. No defensiveness, no JADE. No emotional content. Polite but closed.**]

The BIFF response is designed for written communications with high-conflict people: email, text, legal documents, social media responses. The structure: **B**rief (one to three sentences), **I**nformative (provides only the necessary factual content), **F**riendly (neutral to warm tone without emotional escalation), **F**irm (closes the topic without inviting further engagement).

The critical prohibition is JADE: Justify, Argue, Defend, Explain. Any of these — a defense of your position, an explanation of your reasoning, an argument for your interpretation — opens a loop that the high-conflict person will use to extend the conflict. They are not looking for information; they are looking for material. The BIFF response provides no material.

Example: if someone sends an escalating email attacking a decision you made, the BIFF response provides the factual information directly asked for ("The meeting is on Thursday at 2pm"), expresses a minimal form of goodwill ("I hope that works for you"), and provides no opening for continuation ("Please let me know if you have questions about logistics").

Failure mode: Any emotional content — defense, apology beyond the minimal, explanation of reasoning — opens the JADE loop. The high-conflict person will use the defense to extend the argument, the apology to extract more concessions, and the explanation to find new grounds for objection. Failure reveals engagement on the high-conflict person's terms rather than yours.

Hard case applicability: Specifically designed for high-conflict personality encounters. Less applicable in other contexts.

6.3 MOVE 4.3 — THE TEN-YEAR BELIEF REVISION TEMPLATE

Source: Boghossian & Lindsay (2019, §0.65.15–§0.65.21)
Native type: Persuasion → inquiry; for maximum-confidence cases

> **"Think back ten years ago. Have any of your beliefs changed since then?"** *(If yes:)* **"Ten years ago you believed something you don't believe now. How do you know this belief won't be one of the ones you'll look back on and revise?"**

For someone at maximum confidence who has rejected the disconfirmation question, this move uses their own history of belief change to open the disconfirmation door indirectly. Rather than

asking "could you be wrong about this?" — which the person has already declined — it asks "have you been wrong about other things?" The answer to the second question is almost always yes, and the yes creates space for the first question to be revisited.

The face-saving structure of this move is its primary virtue. It does not demand that the person be wrong now — it only asks them to acknowledge the general human pattern of belief revision over time. They can agree to the general pattern without committing to any revision of their current position. That acknowledgment, however, is the opening: it establishes that they are the kind of person who has updated before, which makes future updating less threatening to their identity as a coherent epistemic agent.

Failure mode: The move fails if the person reports having essentially never changed any beliefs — which can happen with high-conflict personalities or with people whose identity is heavily invested in consistency. In these cases, the move can be gently tested: "Even on small things — preferences, interpretations of events?" Most people will find at least one example if prompted.

Hard case applicability: Anti-intellectualism and narrative-pushing. Less applicable to high-conflict personalities.

6.4 MOVE 4.4 — DISAVOWING YOUR OWN SIDE'S EXTREMISTS

Source: Boghossian & Lindsay (2019, §0.27.2–§0.27.4)

Native type: Persuasion, cross-partisan contexts

> **"I think [specific extreme position on your side] goes too far when [specific behavior]. That's not my position."**

Voluntarily naming how your own side goes too far creates immediate moral common ground, signals that your intellectual commitments are not identical to your tribal identity, and distinguishes you from the caricature of "people on your side" that the other person may be anticipating. It is a status move in Storr's framework: you are demonstrating that you are not a tribal zealot, which lowers the status threat that your position otherwise represents.

The move must be specific. A generic "there are extremists on both sides" is recognized immediately as diplomatic boilerplate and accomplishes nothing. Naming a specific position or behavior that you think goes too far, and explaining why you think so, is evidence that the disavowal is genuine.

Failure mode: If the disavowal is instrumental — chosen because it will work rather than because it is genuinely believed — the inauthenticity is likely to show. The move also fails if the disavowal is so extensive that it makes your actual position unclear. Failure reveals a performance of intellectual independence rather than the thing itself.

Hard case applicability: Primarily applicable to anti-intellectualism encounters, where the person often has a strong sense that expertise-defenders are uniformly arrogant and partisan. Moderate applicability to narrative- pushing.

7

5. DIAGNOSTIC MOVES

Diagnostic moves are used in real time to identify what game you are in and whether the person's epistemic regime allows for joint inquiry. They are not about gathering ammunition — they are about calibrating your approach, specifically about deciding how much of the playbook is available to you in this conversation and what realistic success looks like.

The most important calibration question is this: is joint inquiry achievable, or is the realistic goal something narrower — containing the damage, de-escalating, graceful exit, or protecting third parties who are watching? The diagnostic moves exist to answer this question with some rigor rather than defaulting to optimism or pessimism.

7.1 MOVE 5.1 — THE COMPOSITE DIAGNOSTIC

Agent's construction, synthesized from Boghossian & Lindsay (2019), Eddy (2018), Walton (2013), Galef (2021), and Rauch (2021)

This is not a tested instrument. It is a heuristic synthesized from five sources, offered as a working tool pending empirical validation. Use it as a checklist, not a scoring system — two or more red flags suggest that the realistic goal is narrower than joint inquiry.

> **Test 1 (Disconfirmation):** "Under what conditions could you be wrong about this?" — "None" or implausibly remote conditions = red. **Test 2 (Proportionality):** Are emotional reactions dramatically disproportionate to the conversational stakes? Extreme all-or-nothing statements? = red. **Test 3 (Tracking):** Is the person engaging with what you are actually saying, or pushing their own point regardless of your responses? = red if not tracking. **Test 4 (History):** "Can you think of an area where you've updated a belief recently?" — "Nowhere" or resistance to the question = red. **Test 5 (Falsifiability):** Are the person's claims in principle falsifiable — could evidence count against them? = red if not.

Two or more red lights do not mean the conversation is over. They mean the goal has changed. With two red flags, the realistic goal shifts from "convert to joint inquiry" to "plant doubt that may germinate later." With four or five red flags, the realistic goal shifts further: exit with dignity, protect the audience if one exists, do not invest further.

7.2 MOVE 5.2 — GAME-RECOGNITION QUESTIONS

Source: Walton, Methods of Argumentation (2013)
Native type: All types, as meta-level diagnostic

> **"What would it take for you to change your view on this?" "What are we trying to accomplish in this conversation?" "Are we trying to figure out what's true, or are you trying to persuade me of something specific?"**

These are process questions rather than content questions. They reveal which constitutive rules the other person is operating under — specifically, whether they have any truth-tracking rule in

operation at all. Someone conducting genuine inquiry will engage with the first question seriously and thoughtfully. Someone in eristic mode will find it strange or threatening: the question implies that there is something to figure out together, when they believe they already know what needs to be established.

The third question is blunt and should be used late in a conversation, after other diagnostic indicators have suggested eristic mode, rather than as an opening. Used early, it can sound adversarial; used late, it names a pattern that has already become visible.

Failure mode: If you ask these questions in a tone of accusation rather than genuine curiosity, they function as aggressive moves rather than diagnostic ones. The questions are only useful when the practitioner would genuinely update based on the answer.

Hard case applicability: All three, but particularly applicable as a gateway test before investing significant effort in any of the other techniques.

7.3 MOVE 5.3 — THE WEB METHOD

Source: Eddy, 5 Types of People Who Can Ruin Your Life (2018)
Native type: HCP diagnostic; pre-conversation assessment

> **W — Words:** Is the person making extreme all-or-nothing statements? ("You always," "Everyone knows," "No one could possibly") **E — Emotions:** Are emotional reactions dramatically disproportionate to what would be expected given the actual stakes? **B — Behavior:** Is there a pattern of escalating behavior regardless of your responses — behavior that does not change when you accommodate, explain, or concede?

Eddy's WEB method is an assessment tool, not a conversational move. It is most useful before deciding how much to invest in a conversation, or during a pause (Move 3.2, the balcony) when you are evaluating what is happening.

The key diagnostic is the B — behavior patterns over time. Extreme words and disproportionate emotions can appear in any heated conversation; they are not specific to high-conflict personalities. What distinguishes HCP encounters is the behavior pattern: escalation that continues regardless of what you do. If accommodation produces gratitude but no de-escalation, or if concession produces demand for further concession, you are in HCP territory.

Eddy's practical guidance: high-conflict people cannot be reasoned out of positions they were not reasoned into. The intervention is not argument but structure — clear limits on acceptable behavior, consequences for violations, and disengagement from emotional bait.

Failure mode: The WEB method can be misapplied to dismiss people who are legitimately upset. High emotion is not diagnostic of HCP — it is the behavioral pattern that matters. Failure reveals that the practitioner is using the diagnostic to avoid engagement rather than to calibrate it.

Hard case applicability: Specifically designed for high-conflict personality encounters.

7.4 MOVE 5.4 — THE SCHOPENHAUER CATALOG (DIAGNOSTIC USE ONLY)

Source: Schopenhauer, The Art of Always Being Right (1831)

Use as a diagnostic for recognition, not as a playbook for deployment.

Schopenhauer's 38 eristic stratagems are a catalog of bad-faith moves — a taxonomy of what someone does when they are trying to win an argument they cannot win honestly. The catalog has no peer-reviewed validation and is not a scientific instrument. Its value is practical recognition:

when something feels wrong in an argument, Schopenhauer's list often names what is happening. The most important for real-time detection:

Extension (Stratagem 1): Your argument is extended beyond its intended scope to make it look absurd. You said X in context C; the extension presents X as if you said it universally.

> **Recognition script:** "That's an extension of what I said. My actual claim was more limited — specifically, [restate original claim]."

Homonymy (Stratagem 2): A word is used in two different senses within the same argument, exploiting the ambiguity to attack a position you didn't take.

> **Recognition script:** "I think we may be using [word] differently. When I said it, I meant [sense A]. If you're using it in the sense of [sense B], we may actually agree."

Generalization (Stratagem 3): A specific, qualified claim is converted into a universal statement that is easier to refute.

> **Recognition script:** "I said [specific, qualified claim] — I didn't say [universal version]."

Appeal to authority (Stratagem 30): A credentialed assertion substitutes for argument, especially when the authority cited is loosely relevant.

> **Recognition script:** "That's [Authority's] view. What's the argument behind it?"

Personal attack (Stratagem 38): When all else fails, the argument shifts from the claim to the person making it.

> **Recognition script:** Stay on the argument. "I'm happy to discuss whether my reasoning is right — can we look at [specific step]?"

The practitioner's response to recognized stratagems should generally be naming, not matching. Meeting extension with extension is eristic; naming the move and returning to the original claim is inquiry.

8

6. DISENGAGEMENT MOVES

Knowing when to stop is as important as knowing how to continue. Boghossian and Lindsay (2019) are unusually honest about this: "The most powerful defense against any intervention is refusing to engage" (§0.88.2). No technique works on someone who has closed off. This is not a failure of technique — it is a feature of autonomous cognition. People have the right not to change their minds, and a conversation that tries to force belief change through exhaustion is no longer serving inquiry.

The realistic goal shifts across a spectrum as a conversation reveals itself. Joint inquiry is the ideal; seed-planting — leaving a residue of doubt that may germinate after the conversation is over — is a realistic fallback. Boghossian notes that many belief changes reported by participants in Street Epistemology conversations happened hours or days after the conversation ended, not during it (§0.94.3). The conversation's effects are not always visible in the conversation. When to stop is therefore in part a question of when continued engagement will add to the seed and when it will merely confirm the person's existing reasons for resistance.

8.1 MOVE 6.1 — PLANTING AND RETREATING

Source: Boghossian & Lindsay (2019, §0.95.6, §0.97.4)
Native type: All types, as closing move

> **"I appreciate you thinking through this with me. I'll let you sit with those questions — I find they're often more interesting to think about on your own."**

When doubt has been planted — when the person has expressed any visible uncertainty, asked a question they hadn't asked before, or shown interest in a distinction they hadn't previously considered — do not press the advantage. The optimal moment to exit is when the other person first shows uncertainty. Pressing at that moment re-triggers defensiveness: the person closes off the doubt by finding new reasons to resist.

Instead, acknowledge their engagement with warmth, thank them for the conversation, and leave the doubt to develop independently. You are planting a seed, not extracting a commitment. Seeds need time and space; they do not need more water immediately.

Tavris and Aronson's cognitive dissonance research explains why restraint matters. The dissonance between a newly acknowledged doubt and an existing commitment needs time to resolve. The self-justification engine is slower than the dissonance — if you return too quickly with more pressure, the engine activates and forecloses the doubt before it can do its work.

Failure mode: Pressing when the person shows doubt — adding more evidence, making a stronger argument, asking another probing question — is the single most common error after good work has been done. Failure reveals the practitioner's righting reflex: the compulsion to close the sale while the person is listening. The compulsion is destructive precisely when it feels most justified.

Hard case applicability: All three, but particularly important in anti-intellectualism encounters, where any sense of being pressured will activate the identity defenses that the whole conversation has been working to circumvent.

8.2 MOVE 6.2 — GRACEFUL EXIT

Source: Boghossian & Lindsay (2019, §0.17.2–§0.17.8)
Native type: All types, terminal

> **"I've really enjoyed thinking through this with you. I don't think we're going to resolve it today, but I've learned something from this conversation." "I can see we see this very differently. I'm going to keep thinking about what you've said."**

Exit signals: frustration has become the dominant emotion in the conversation; the other person wants to disengage; doubt has been planted and the person seems on the edge of shifting (the optimal planting moment); repeated disconfirmation attempts have all been dismissed; the composite diagnostic (Move 5.1) shows four or five red flags.

The graceful exit has two components: an acknowledgment of the conversation without claiming to have won it, and an expression of genuine respect for the person (not necessarily their position). The second component is not merely polite — it is strategically correct. The conversation does not end when you stop talking; it continues in the other person's mind. How you leave shapes the emotional valence of what follows. Leaving with contempt, even implicit contempt, closes the door. Leaving with genuine respect — "I'm going to keep thinking about what you've said" — is an invitation to the other person to do the same.

Never express contempt, even when it is genuinely felt. The audience — if any exists — sees how you leave. The other person remembers how they were treated.

Failure mode: The graceful exit fails when it is used as avoidance — when you leave before doing any real work. It is a closing move, not an alternative to the playbook. Failure reveals that the practitioner wanted to disengage without wanting to have engaged.

Hard case applicability: All three. The exit protocols differ by context: from high-conflict personalities, exit without providing material; from anti-intellectualists, exit while leaving respect intact; from narrative- pushers, exit while acknowledging what was valuable in what they were trying to say.

8.3 MOVE 6.3 — PROTECTING THE AUDIENCE

Source: Rauch, The Constitution of Knowledge (2021)
Native type: Public or semi-public eristic encounters

> **Address your argument to the room, not to the hostile interlocutor. State clearly what they said, what it assumes, and why you think the reasoning doesn't hold — calmly and without contempt. Then disengage.**

Rauch's insight: in public discourse, the goal of honest interlocutors is not to convert the troll or ideologue but to maintain epistemic standards in front of people who are observing. The audience is often the primary constituency; the interlocutor is secondary. The person you are talking to may be performing for their audience while you are talking to yours. Recognizing this changes the goal entirely.

When the interlocutor cannot be reached — the composite diagnostic (Move 5.1) shows multiple red flags, the conversation has become visibly eristic — address the substance of the argument to the broader audience. Acknowledge what was said, state why the reasoning is problematic, do so without contempt or condescension, and decline to continue the exchange with the interlocutor.

The critical condition: Rauch's move is genuinely a public epistemic service only when it is conducted with the standards of the liberal epistemic regime he defends — evidence, argument, good faith, openness to correction. If it degenerates into point-scoring for one's own audience, it has become a mirror image of the bad-faith discourse it is supposed to counteract.

Failure mode: The move fails when "protecting the audience" becomes a rationale for rhetorical performance — for scoring points with your own tribe rather than modeling good epistemic practice for a genuinely mixed audience. Failure reveals that the practitioner has drifted into eristic mode while believing they are conducting inquiry.

Hard case applicability: Primarily relevant to anti-intellectualism in public or semi-public contexts — debates, comment sections, public forums — where the immediate interlocutor is unlikely to update but where the practice of honest argument is itself worth demonstrating.

8.4 A NOTE ON REALISTIC GOALS

The moves in this section, and in the playbook as a whole, operate against a baseline expectation that needs to be stated clearly: most difficult conversations do not end with the other person changing their mind. The research literature on belief change — from Motivational Interviewing's clinical trials to Broockman and Kalla's deep canvassing studies — shows meaningful effect sizes compared to control conditions, but effect sizes describe population averages. Any given conversation with any given person may produce no visible change.

What changes in the aggregate, across many conversations over time, is different from what happens in any single exchange. The practitioner who understands this can maintain the scout posture — genuinely curious, not invested in any particular outcome — rather than treating each conversation as a test of whether the techniques work. Techniques work probabilistically. Conversations are individual events. The goal of the playbook is to improve the probability of productive outcomes, not to guarantee them.

Boghossian and Lindsay close their practical guide with a passage that serves well as the close of this section: the purpose of good-faith dialogue is not to change people's minds but to make mind-changing possible. You are not the agent of the change. You are the person who tries to create the conditions under which change can happen — conditions of safety, genuine curiosity, face-saving exits, and honest inquiry. What happens after that is, properly, up to the other person.

9

PART III: HARD CASES

The three hard cases in this review form a gradient of convertibility — a spectrum from conversations that are fundamentally rescuable to conversations where the goal of joint inquiry must be abandoned entirely in favor of narrower and more defensive aims.

It is worth being explicit about this gradient at the outset, because the most common failure mode when entering a difficult conversation is misidentifying which hard case you are in. Treating a narrative-pusher as if they were a high-conflict personality produces needless antagonism toward someone who genuinely needs help articulating their position — it is, in van Eemeren's terms, a Rule 5 violation in reverse: you are attributing to them a worse position than the one they actually hold, and then responding to the attribution. Treating a narcissistic interlocutor as if they were a narrative-pusher — applying patient Socratic questioning to someone whose conversational agenda is not epistemic at all — produces extended exposure to behavior that harms you and reinforces the pattern. The diagnostic moves from the playbook (Section 5) exist to prevent this misidentification. Use them before committing to a strategy.

The gradient runs like this. Narrative-pushing — the first case — is the most convertible: the person has a real intuition they haven't articulated, and careful questioning can help them find it. Anti-intellectualism — the second case — is partially convertible: the person is resistant to evidence and expertise, but their resistance has epistemic roots (if distorted ones), and there are techniques with demonstrated efficacy. Narcissistic and high-conflict personalities — the third case — are usually not convertible: the conversational dysfunction is structural, rooted in personality organization rather than epistemics, and the realistic aims narrow to containment, de-escalation, and exit.

For each case, this section describes what it looks like, what is happening underneath the surface behavior, which techniques apply, what a realistic successful outcome looks like, and when to walk away. It also describes something the original playbook does not adequately address: what it feels like from your side, and how your own cognitive and emotional responses can either help or sabotage the interaction.

9.1 HARD CASE 1: NARRATIVE-PUSHING

9.1.1 *What It Looks Like*

The narrative-pusher is not trying to deceive you. They have an intuition — often a genuine and important one — that they have not been able to articulate clearly. They know the conclusion; they cannot state the premises. They know they are bothered by something; they cannot say what. In Walton's framework, they appear to be in a persuasion dialogue, but their commitment set is structurally underspecified: they can't precisely state what they're committed to, so they can't be precisely committed to anything.

The straw man danger here is severe — not primarily because they will strawman you (they might), but because you will inadvertently strawman them. You hear a vague claim, you assign it the nearest crisp formulation, and you engage with the crisp formulation. The problem is that the crisp formulation is yours, not theirs. You are now debating a position that nobody in the room

actually holds, while the person across from you feels increasingly frustrated that you keep "not getting it." They are right. You are not getting it — not because you failed to listen, but because they haven't yet said what they mean.

This is the PI's original framing — "interlocutors who have a point to make but don't fully understand it" — and it does not map cleanly onto any single source's category. It is close to what Boghossian and Lindsay (2019) call "the ideologue case," but distinguishable: the ideologue has a clear (if unfalsifiable) position, defended with commitment. The narrative-pusher doesn't have a clear position yet. They are pre-ideological in a certain sense — the position is still forming, and the conversation is happening at the moment of formation. This makes them more open to being helped than an ideologue, but it also makes them more emotionally vulnerable to feeling misunderstood, because they are right that they haven't been understood — not because you have failed to listen, but because they haven't yet said what they mean.

9.1.2 *The Observable Markers*

The narrative-pusher changes their stated position under questioning without acknowledging the change — not out of bad faith, but because each new formulation feels like a closer approximation to what they mean. They get more emotional when pressed for specificity, not less — the emotional intensity reflects the gap between what they feel and what they can say. They retreat to vaguer formulations when offered crisper ones: "That's sort of it, but not exactly." They say "you know what I mean" or "I can't explain it but it's real." They defend the conclusion vigorously while being unable to articulate the premises.

These are not signs of dishonesty. They are signs of a person trying to communicate something they haven't yet found words for. The frustration is real and legitimate — the experience of knowing something you cannot say is one of the most cognitively uncomfortable states there is, and being pressed to say it more clearly when you can't feels like being tested on material you haven't studied.

9.1.3 *What It Feels Like From Your Side*

Here is where the practitioner must be honest with themselves. Talking to a narrative-pusher is annoying. It is annoying because you are doing the intellectual work of steelmanning their position, offering precise formulations, asking clarifying questions — and they keep saying "no, that's not quite it" without being able to tell you what it is. The temptation is to conclude that they don't have a point at all — that the vagueness is the point, that they are just complaining or emoting or performing a position they absorbed from somewhere.

Sometimes that assessment is correct. But not as often as you think. Haidt's elephant-and-rider metaphor (2012) is directly relevant here: moral intuitions arrive first, and the conscious reasoning mind constructs justifications after the fact. The narrative-pusher's intuition may have arrived via the elephant — a felt sense that something is wrong, shaped by experience, emotion, and moral foundations that they may not be able to articulate in the language of the rider. Your questions are addressed to the rider. The rider doesn't have the answer yet. The elephant does, but elephants don't speak in propositions.

The practitioner's internal danger is impatience that curdles into contempt. The moment you decide that the other person is fundamentally confused — that they simply don't have a point — you stop listening for the point, and you will miss it if it arrives. Bohm's analysis applies: your opinion that "this person doesn't have a real argument" has become something you are defending, and the defensive posture makes you unable to hear the argument when it finally surfaces.

9.1.4 *What Techniques Apply*

The narrative-pusher actually needs Socratic questioning more than any other interlocutor in this hard-cases section. They are genuinely confused, and careful questions can help them discover what they actually think. This is the hard case where conversion to joint inquiry is most consistently achievable — not because you will persuade them of your view, but because you can help them arrive at a clear version of their own.

Humble Inquiry (Schein, 2013) is the primary orientation: draw out what they actually mean rather than what they said. Ask for examples, stories, specific cases rather than principles. "Can you give me an example of what you're describing?" does more work here than any argument. The example forces specificity without demanding it — the person produces the concrete case themselves and can recognize whether it captures what they meant.

The reason this works where direct questions fail is precisely Schein's distinction between humble inquiry and confrontive inquiry. "What exactly is your position?" is confrontive — it demands an answer the person doesn't have, which produces either a defensive lunge toward a position they don't actually hold or an admission of confusion that feels like defeat. "Can you think of a time when this happened?" is humble — it asks for something they do have (an experience) rather than something they don't (a thesis).

Commitment Tracking (Walton, 2013) shifts from gotcha to genuine aid when done with curiosity. "Help me make sure I understand your position. Are you saying [formulation A] or [formulation B]?" is not a trap — it is a genuine service. Done with an adversarial frame, the question reveals inconsistency; done with a collaborative frame, the question helps the person triangulate toward what they actually believe. The difference is not technique but disposition. The same words can be humble inquiry or prosecutorial cross-examination. The interlocutor almost always knows which one they're getting.

Rapoport's Rules, applied generously, create the conditions for the narrative-pusher to correct you. Before any engagement with the substance of what they've said, restate their position more clearly than they stated it. If you've understood correctly, they will recognize themselves. If you've introduced a distortion — even inadvertently, even in the direction of making their position more coherent than it was — they will be prompted to correct you toward a more accurate formulation. This is the productive move. Letting them correct your restatement is often more useful than any question you could ask. Their corrections reveal the shape of what they mean more precisely than their original statements did.

Narrative Elicitation, drawn from the deep canvassing protocol (Broockman & Kalla, 2016), asks for a personal story rather than an abstract position. "Can you tell me about a time when you experienced something like this?" The personal story reveals the lived experience behind the intuition and gives you something concrete to engage with rather than a vague principle to argue about. This is not a rhetorical technique deployed to soften them up. It is a genuine request to understand where the intuition comes from — and the answer to that question usually reveals far more about what they actually mean than any amount of propositional clarification.

The Broockman and Kalla finding is worth invoking precisely here: in the deep canvassing protocol, the mechanism that produced durable attitude change was narrative exchange — the sharing of personal stories — not argument. When the narrative component was removed and only the argumentative content preserved, the effect disappeared. For the narrative-pusher, narrative elicitation is the primary technique not because it will change their mind but because it surfaces what they are actually saying. The story they tell you will often contain the thesis they couldn't articulate in the abstract.

A concrete scenario. Someone says: "I just think schools are failing kids these days." You ask what they mean. They gesture vaguely at declining standards, kids on phones, teachers who don't care. You could engage with any of these claims factually — test scores, screen time research, teacher workload data. But the factual engagement will miss the point because the stated claims are not the point. The point is somewhere underneath them.

So instead: "What's your experience been? Did something happen with your kid, or with a school you know?" The answer might be: "My nephew's teacher told his parents he was fine, and then he got to high school and couldn't read at grade level." Now you have something real. The intuition — "schools are failing kids" — was tracking a specific failure of institutional honesty. The person's thesis, once you help them find it, might be something like: "The system has incentives to pass kids through rather than admit they're struggling." That is a real claim with real evidence behind it, and it is a claim you might actually agree with. But you would never have found it by arguing about whether schools are failing in the aggregate.

9.1.5 *The Realistic Goal*

Joint inquiry is often achievable with the narrative-pusher. The conversation's realistic successful outcome is a person who has, through the process of the conversation, arrived at a clearer and more articulable version of their position — even if that position disagrees with yours. This is the one hard case where "helping them find their argument" is the correct goal. It is also the one case where you might discover that their intuition was tracking something real that you had missed.

The success marker is not agreement. It is precision. The conversation has succeeded when the person can say "what I mean is X" with enough specificity that X can be engaged with, evaluated, agreed with, or contested. Getting them to that point is an act of intellectual generosity, not a conversational strategy.

There is a secondary success that is worth naming because practitioners often don't notice it: the relationship improves. A person who feels that you genuinely helped them figure out what they think — rather than defeating their argument or dismissing their confusion — will seek you out for future conversations. You become, in their experience, a person worth talking to. This is not a trivial outcome. Over time, being the person others seek out when they're trying to think through something difficult is among the most valuable social positions one can occupy.

9.1.6 *When to Walk Away*

When genuine curiosity has been fully exhausted and the person becomes visibly frustrated with the process of clarification itself — when careful questions are experienced as attacks, when the request for a specific example is heard as mockery — the conversation has shifted into a different category. At that point, the confusion has become defensive. The person is no longer struggling to articulate something; they are protecting something.

The diagnostic: did the emotional tone shift from "I'm frustrated because I can't find the words" to "I'm angry because you keep challenging me"? The first is productive frustration — the person is working. The second is defensive frustration — the person has stopped working and started fighting. The difference is often audible in a single sentence.

The appropriate response is not to push harder but to exit cleanly: "I haven't quite grasped what you're driving at — I'll keep thinking about it." This leaves the door open for a future conversation in a lower-stakes context, and it does not provoke the defensive escalation that continued questioning would produce. The invitation is genuine — and sometimes the person will come back later with

the words they couldn't find, because the conversation gave them the shape even though it didn't give them the formulation.

9.2 HARD CASE 2: ANTI-INTELLECTUALISM

9.2.1 *What It Looks Like*

Jonathan Rauch (*The Constitution of Knowledge*, 2021) draws a distinction that is essential for understanding this hard case. An **epistemic bubble** is a community or media environment from which contrary evidence has been filtered out — the person hasn't been exposed to good arguments, and exposure with rapport is often sufficient to shift their position. An **echo chamber** (Nguyen, 2018) is structurally different: the person has been exposed to contrary evidence but has been inoculated against it. Sources outside the community are pre-dismissed as corrupt, biased, or controlled. The correction mechanism has been disabled. This is the anti-intellectual condition in its most challenging form.

The distinction matters for technique selection. The person in an epistemic bubble needs exposure plus relationship — they haven't encountered the evidence, and presenting it with care can work. The person in an echo chamber needs something different, because the problem is not informational but epistemological: they have an alternative epistemology that makes mainstream evidence epistemically invisible.

9.2.2 *The Observable Markers*

The echo-chamber anti-intellectual dismisses expert consensus as conspiracy rather than engaging with specific claims — "they're all in on it" rather than "here's what's wrong with that study." They use "do your own research" as a terminal move rather than an invitation. They apply different epistemological standards to preferred and dispreferred sources: demanding certainty of dispreferred sources while accepting any corroboration for preferred ones. They are fluent in the language of skepticism but apply it only in one direction — they can list reasons to doubt mainstream sources but will not apply the same analysis to their alternative sources.

But there is a subtler marker that is easy to miss if you are looking only for the obvious ones. The anti-intellectual often has **a specific personal experience of institutional failure** that serves as the epistemic anchor for their generalized distrust. They were lied to by a doctor. Their community was harmed by a policy that experts endorsed. Their expertise in their own domain — farming, manufacturing, small business — was dismissed by credentialed people who turned out to be wrong about the specific thing the anti-intellectual knew most about. This experience is real. The overgeneralization from it is the problem, but the experience itself is genuine evidence, and treating it as irrelevant or irrational is both factually wrong and strategically disastrous.

9.2.3 *What It Feels Like From Your Side*

This is the hard case that activates what Miller and Rollnick call the **righting reflex** most powerfully. The righting reflex is the urge to correct — to present the evidence, to explain why the experts are actually right, to make the person see. It is the most natural response in the world, and it is almost always counterproductive.

The righting reflex is counterproductive because it performs exactly the dynamic the anti-intellectual has already learned to resist. They have encountered credentialed people who dismissed

their experience and told them what to think. You are now being another one of those people. From their perspective, you are confirming their thesis: experts dismiss ordinary people's knowledge and experience. The fact that you are right about the specific claim does not matter, because the conversation is not primarily about the specific claim. It is about the epistemological relationship between expertise and lived experience, and on that terrain, you are losing the moment you reach for your credentials or your citations.

Tavris and Aronson's (2007/2020) analysis of cognitive dissonance is the mechanism. The anti-intellectual has made a series of commitments — social, emotional, sometimes financial — based on their alternative epistemology. Each commitment raises the cost of admitting error. They are not defending a single claim. They are defending a structure of commitments that, if any single claim were conceded, would begin to unravel in ways that threaten their self-concept, their community membership, and their sense of having been a reasonable person all along. Tavris and Aronson call this the "pyramid of choice": each small commitment narrows the range of positions that remain psychologically available, until the person stands at the top of a pyramid and cannot step down without falling.

Your impatience — your feeling that "this should be obvious" — is itself a signal that you are not in the right posture. The right posture begins with the recognition that the person's position, while wrong in its conclusions, is not crazy in its origins. They observed something real and drew too broad a conclusion from it. The technique is to engage the observation without endorsing the conclusion.

9.2.4 *What Techniques Apply*

On the limits of fact-correction: Ecker et al. (2022, *Nature Reviews Psychology*) provides the authoritative current synthesis. Corrections of misinformation generally work — the much-cited "backfire effect" (corrections strengthening false beliefs) has not replicated robustly (Wood & Porter, 2019). What IS established is that corrections are often insufficient, particularly when the belief is identity-relevant. The situation is not hopeless, but simple fact-correction is unlikely to be the primary vehicle for change, and deploying it as the primary strategy will usually produce disengagement rather than updating.

Motivational Interviewing applies directly here through its central insight: the righting reflex is the enemy. The MI move is to ask about their reasoning rather than present yours. "What's your sense of what's happening with [topic]?" / "How did you come to see it that way?" These questions are not rhetorical. They are genuine inquiries into the person's epistemic history, and that history often contains experiences — of being lied to by institutions, of having their expertise in their own domain dismissed, of watching credentialed people make catastrophic errors — that have some evidentiary warrant even if the conclusions drawn from them are too broad.

Acknowledging the legitimate kernel without endorsing the overextended conclusion is the most important MI move in this context: "I think you're right that [specific institution] got [specific thing] badly wrong. Where I think it goes further than the evidence supports is [specific claim]." This is not rhetorical flattery. It is honest acknowledgment that anti-intellectual positions are often tracking something real about institutional failure, even when they draw the wrong epistemological conclusions from it.

The honesty is critical, and it requires that you actually mean it. Schein's insight about performed curiosity applies: the anti-intellectual has been on the receiving end of condescension from experts for long enough to detect it instantly. If you are pretending to take their experience seriously as a technique to gain rapport, they will know, and the pretense will confirm their model of how

experts operate. The acknowledgment only works if it is genuine — if you actually recognize that their distrust of institutions has some warrant, even though you disagree with the conclusions they've drawn.

Technique Rebuttal (Schmid & Betsch, 2019, *PNAS*) is the most directly applicable experimental finding in this case. Schmid and Betsch found that addressing the *technique* of misinformation — the rhetorical strategies used to manufacture doubt — was more effective than addressing the specific content of false claims. The practical move: discuss the epistemological process rather than the specific claims.

"How do you evaluate which sources to trust?" is a question that opens epistemological reflection without attacking any specific belief. The person has to think about their own method, which is cognitively different from defending a conclusion. "What would count as good evidence for or against this position?" asks them to specify their epistemic standard — and the act of specification often reveals that the standard is inconsistent. "What would it take to change your mind?" is the Street Epistemology disconfirmation question applied directly. These questions engage the mechanism of belief formation rather than the content of the belief.

A concrete scenario. Someone says: "I don't trust the CDC. They've been wrong about everything." You feel the righting reflex. You want to list the things the CDC got right. Don't.

Instead: "What specifically did they get wrong that sticks with you the most?" The answer will usually be specific — the shifting mask guidance early in COVID, perhaps, or a recommendation that was later reversed. These are real things that happened. Acknowledge them: "Yeah, the mask guidance was genuinely confusing. They changed the recommendation and didn't explain why clearly enough."

Then: "When you're deciding what to trust about a health question, what do you look at?" This is the epistemological question. The person will describe their process. It might be "I talk to my doctor" or "I read studies myself" or "I trust people who have actually been through it." Each of these is an epistemological claim about what constitutes good evidence, and each can be engaged on its own terms.

What you are doing — and this is the structural move that Schmid and Betsch validated — is shifting the conversation from "is the CDC right about X?" (topic rebuttal) to "how do you and I figure out what to believe about health questions?" (technique rebuttal). The second conversation is one where you are genuine collaborators rather than adversaries. You both have the problem of figuring out what to trust. You both have experiences of experts being wrong. You both need a method. That shared problem is the ground on which joint inquiry becomes possible.

Street Epistemology's disconfirmation sequence follows naturally: "Under what conditions would you update your view that the CDC can't be trusted?" If the answer is "none" or requires conditions that are definitionally impossible ("if they admitted they were lying"), the belief is in non-disconfirmable territory. This is not an accusation — it is information. Follow the Boghossian sequence: "What I'm noticing is that no evidence you've described could change your view — that seems like a problem for any belief, mine included. What do you think makes a good reason to update?"

The "mine included" is load-bearing. It signals that you are applying the same standard to yourself — that disconfirmability is a universal epistemic virtue, not a weapon you deploy selectively against beliefs you disagree with. If the person asks you the same question — "What would change your mind about the CDC?" — you should have an honest answer. If you don't, you are doing exactly what you're asking them not to do.

Inoculation/Prebunking (Cook et al., 2017; van der Linden, 2022) is the most empirically supported proactive intervention, but its applicability to existing anti-intellectual positions is limited. Inoculation works by exposing people to weakened forms of misinformation techniques

— with explanations of the manipulation — before they encounter full-strength versions. It is a proactive tool, most useful before the echo chamber has fully formed. For someone already fully in the echo chamber, prebunking is of limited use in real-time, though it remains valuable as a population-level public health intervention.

9.2.5 *The Realistic Goal*

For the person in an epistemic bubble — not yet in an echo chamber — the realistic goal is genuine updating, and it is often achievable with the right combination of rapport and evidence.

For the echo-chamber anti-intellectual, the realistic goal narrows. Rauch's description of the "epistemic sociopath" — someone whose identity investment in anti-expertise is so total that no correction can reach them — names the outer limit. For most people short of that extreme, the realistic goal is not agreement but keeping the question open: maintaining the possibility that the person might, later, in a different context, reconsider. Planting doubt rather than achieving conversion. Demonstrating that someone who accepts mainstream expertise can take their concerns seriously without being stupid or corrupt.

This is a modest goal, and it should be stated modestly. But it is not a trivial one. The anti-intellectual's model of the world includes the premise that people who trust experts are either naive or tribal. Every encounter with someone who trusts experts *and* genuinely listens *and* acknowledges institutional failures *and* admits uncertainty is an encounter that doesn't fit the model. One encounter won't break the model. Ten might not. But the model requires maintenance — it requires every encounter to confirm it — and the encounter that didn't confirm it is the one that sits at the edge of awareness, generating the low-grade cognitive dissonance that is, over time, the precondition for updating.

In public conversations, the audience is often the real target even when the nominal interlocutor is unreachable. A composed, respectful engagement with an anti-intellectual interlocutor — one that takes their concerns seriously while declining to accept their epistemology — communicates to everyone watching what good-faith epistemic practice looks like. This is not cynical audience-management. It is the recognition that most people in any audience are in the epistemic bubble, not the echo chamber. They haven't been inoculated against your evidence. They are watching to see whether the expert-trusting person is the reasonable, curious one — or the condescending, dismissive one. Your behavior in the conversation determines which model they update toward.

9.2.6 *When to Walk Away*

When the disconfirmation test produces "nothing could change my mind" AND the epistemological question produces a similar response — when the person's self-concept is entirely bound up in epistemic resistance — you are no longer in a dialogue about the topic. You are in a performance.

Rauch's term for the adversarial strategy of some anti-intellectuals is useful here: some are not confused about the facts but are actively trying to destabilize shared epistemic standards as a tactic. For these interlocutors, continued engagement on the merits gives the performance legitimacy it does not deserve.

But be slow to conclude this. The temptation to categorize someone as an irredeemable epistemic sociopath is itself a form of the fundamental attribution error — explaining their behavior by their character rather than by their situation. Most anti-intellectuals are not performing. They are genuinely operating within an alternative epistemological framework that produces different conclusions from the same evidence. The framework is wrong, but the person inhabiting it is

not acting in bad faith. They believe what they believe for reasons that make sense within their framework, and the fact that their framework is defective does not make them defective.

The appropriate exit: "I think we have fundamentally different standards for what counts as good evidence here. That's a real disagreement and I don't think we're going to resolve it today. But I'd encourage anyone listening to look at [specific source] if you want to see the evidence I find most compelling." Address the last sentence to the audience, not the interlocutor. Then stop. Do not re-enter. The conversational exit is complete when you have given the audience something and given the interlocutor nothing further to escalate against.

9.3 HARD CASE 3: NARCISSISTIC AND HIGH-CONFLICT PERSONALITIES

9.3.1 *What It Looks Like*

The key diagnostic distinction for this hard case is this: in Cases 1 and 2, the conversational dysfunction is epistemic. The person is confused, or defended against evidence, or committed to an alternative epistemology — but their conversational behavior is oriented, however distortedly, toward some relationship with reality. In Hard Case 3, the dysfunction is structural. It is driven by personality organization rather than by what is being argued about. The topic is an occasion for a different agenda entirely: status assertion, control, grievance expression, identity defense.

Grapsas et al. (2019, *Psychological Bulletin*, approximately 290 citations) provide the most rigorous academic account of the underlying mechanism. Narcissism operates as a status-pursuit mechanism with two distinct motivational poles: admiration-seeking (the approach motivation) and rivalry (the avoidance motivation activated when status is threatened). The conversational implication is decisive. Initial engagement with a narcissistic interlocutor may look like genuine interest — warm, curious, attentive. But this engagement is conditional on being affirmed. Any challenge — even a gentle factual question — can trigger the switch from admiration-seeking to rivalry, at which point the person becomes adversarial in a way that bears no relationship to the conversational content. What happened is not that you made a bad argument. What happened is that the status dynamic shifted.

Bill Eddy's clinical framework (*5 Types of People Who Can Ruin Your Life*, 2018; *BIFF*, 2014) extends the analysis from narcissism to a broader category of high-conflict personalities (HCPs): narcissistic, paranoid, antisocial, borderline, and histrionic. For this review's purposes, the common structural pattern across types matters more than the specific diagnosis: all-or-nothing thinking; unmanaged emotional responses; extreme behavior disproportionate to the apparent stakes; and the inability to experience dispute without experiencing it as attack. Will Storr's account (*The Status Game*, 2021) complements Eddy's by explaining the virtue-game variant: some HCPs are not primarily playing a dominance game but a virtue game — they are the righteous one; you are the villain; any challenge to their position is an attack on their moral identity.

9.3.2 *What It Feels Like From Your Side*

This is where the review must be most careful, because this is the hard case where the practitioner's own psychology is most directly at risk.

The initial phase of engagement with an HCP is often disarmingly pleasant. They are interested, attentive, apparently receptive. If you have been working through Hard Cases 1 and 2 with moderate success — helping narrative-pushers find their arguments, gaining rapport with anti-intellectuals

— you may enter this conversation with confidence that your techniques will work here too. They will not.

The moment the dynamic shifts — and it can shift on a single question, a single look, a single implication that the HCP's position might not be entirely correct — the conversation changes character entirely. The warmth disappears. The attacks become personal. The scope expands from the specific topic to your character, your motives, your intelligence, your worth as a person. The escalation is fast, disproportionate, and — this is the distinctive feature — it does not track anything you did.

The experience from the inside is bewildering. You said something entirely reasonable, something that would have been received as a thoughtful question in any other conversation, and the person responded as if you had insulted their mother. Your cognitive instinct is to search for what you did wrong. This instinct is a trap. You did nothing wrong. The dysregulated response is not a reaction to your conversational move; it is a reaction to a status threat that your conversational move represented. The HCP is not responding to you. They are responding to the category of experience you triggered: being challenged, being questioned, not being immediately affirmed.

The practitioner's distinctive psychological risk in this case is **self-doubt followed by over-engagement**. The bewilderment of the first shift creates a desire to fix it — to find the right words, to demonstrate good faith, to prove that you weren't attacking. This desire plays directly into the HCP dynamic. Each attempt to clarify, to justify, to defend your intention gives the HCP new material. Eddy's term for this is the JADE trap: Justify, Argue, Defend, Explain. Every JADE response extends the conflict. The impulse to JADE is the righting reflex in a new key — the urge not to correct a factual error but to correct the other person's perception of your character.

The second psychological risk is **contamination of your assessment of the first two hard cases**. After enough encounters with genuine HCPs, you may begin to pattern-match too aggressively — seeing narcissistic dynamics in people who are simply frustrated, confused, or defended. A narrative- pusher who gets emotional when you ask for specifics is not an HCP. An anti-intellectual who gets heated when you question their sources is not an HCP. The HCP pattern is distinctive, and the temptation to apply it too broadly is itself a form of the fundamental attribution error. Reserve this category for its actual instances. They are rarer than you think and more damaging than the other two combined.

9.3.3 *Observable Markers: The WEB Method*

Eddy's WEB method provides the diagnostic:

Words: Extreme all-or-nothing statements — "always," "never," "completely," "totally." Declarations of total certainty combined with extreme moral condemnation of those who disagree. The vocabulary is the first signal.

Emotions: Reactions dramatically disproportionate to the apparent stakes. The emotional intensity is not calibrated to the content of the disagreement — it reflects the underlying status threat. Cycling between sudden warmth (when affirmed) and sudden hostility (when questioned) within a single conversation is a strong marker.

Behavior: Escalation regardless of what you do — conceding makes things worse, not better; agreeing provisionally makes things worse; asking clarifying questions makes things worse. The escalation is not tracking your conversational moves. It is following a pattern that exists independently of you. Checking whether the same patterns appear across the person's relationships is the strongest diagnostic — HCP patterns are typically consistent across contexts, not confined to the current dispute.

A note on the reliability of this diagnostic: the WEB method comes from clinical practice, not from experimental psychology. Eddy's work is practitioner-generated wisdom, not peer-reviewed experimental science. The academic narcissism literature (Grapsas et al., 2019) confirms the underlying mechanism — the admiration-rivalry dual pathway — but the specific diagnostic markers in WEB have not been experimentally validated as a screening instrument. Use WEB as a heuristic, not as a clinical diagnosis. It is useful precisely because it gives the layperson something to look for; it is limited precisely because false positives are possible and the consequences of false-positive labeling are serious.

9.3.4 *What Techniques Apply*

Most conversational techniques have limited or no efficacy with genuine HCPs because they presuppose a conversational partner capable of operating in some good-faith mode — someone who has a truth-tracking orientation, even a distorted one. HCPs often do not. The techniques in the playbook — Socratic questioning, the disconfirmation sequence, narrative elicitation, reflective listening — were designed for partners whose conversational agenda is at least nominally epistemic. When the agenda is status management or grievance expression, these techniques land in the wrong register entirely.

The BIFF Response (Eddy, 2014) is the most useful protocol for written communications with HCPs: Brief, Informative, Friendly, Firm. Answer only what was asked. Provide necessary information. Close without opening further discussion. The critical prohibition is JADE: do not Justify, Argue, Defend, or Explain. Any JADE content gives the HCP material to extend the conflict. The BIFF response removes that material.

The practical difference is striking. An email that explains your reasoning at length — however reasonable the explanation — provides an HCP with ten new attack surfaces. One that states the necessary information briefly and warmly provides almost none. "The meeting is at 3pm. Looking forward to it." is BIFF. "The meeting is at 3pm. I moved it from 2pm because the room was booked, and I wanted to make sure everyone could attend, especially since last time we had scheduling conflicts that I felt were avoidable" is JADE. Every clause after "3pm" is a new line of engagement for the HCP to pull on.

The CARS Method provides the verbal equivalent for real-time encounters:

Connect with empathy: Acknowledge the feeling without endorsing the content. "I can see you feel strongly about this" / "That sounds like it's been really frustrating." This is not capitulation. It is verbal judo — redirecting the emotional force rather than meeting it head-on. George Thompson's insight from Verbal Judo (1993/2013) applies directly: the officer who says "calm down" to an agitated person is asking them to do something they currently cannot do. The officer who says "I can see you're upset" is simply describing what is visible, which the person can accept without losing face.

Analyze options: Indicate briefly what options exist. "There are a few ways we could handle this." This introduces the idea that the conversation can go somewhere without specifying where, which reduces the sense of being trapped.

Respond to misinformation: One clear factual correction only, without elaboration. Not three corrections, not a detailed rebuttal — one statement of fact, stated once, without defensiveness. "Actually, what happened was [X]." Then move on. Repeating the correction or elaborating on it gives the HCP more material and makes you appear defensive.

Set limits: Name what you will do if the behavior continues — not what you will do to them, but what you will do. "If we can't find a way to have this conversation, I'm going to need to step

away." This is a statement of your own behavior, not a threat. It gives them information about consequences without issuing an ultimatum, which would typically escalate.

Storr's Status Framework explains why direct challenges reliably fail with HCPs operating in the virtue-game mode. If their identity in the conversation is "the righteous one who has been wronged," any challenge to their position is a challenge to their moral identity. The playbook move: engage behavior rather than character. "What I'm responding to is [specific thing that happened]" rather than "I think you're being unreasonable." The first is a description of behavior in the world; the second is a judgment of their person, which activates the rivalry response immediately.

The deeper implication of Storr's framework: to de-escalate an HCP in virtue-game mode, it helps to provide some path by which they can retreat from escalation without losing face. This might mean acknowledging a legitimate grievance underneath the disproportionate expression: "I think you're right that [this specific thing] should have happened differently." The concession must be genuine — HCPs are unusually skilled at detecting performed concessions — and it must be specific, not global. "I think you're completely right and I've been unfair to you" is global and will not be believed. "I think you're right that [specific thing] was handled badly" is specific and can be accepted as a partial vindication.

A concrete scenario. You're in a meeting. A colleague responds to your presentation with: "This is exactly the kind of thing that keeps happening in this department. Nobody listens, nobody cares, and then someone comes in and presents this like it's a solution when it's actually the problem." The emotional intensity is disproportionate. The attack is personal but framed as institutional. Your instinct is to defend your presentation — to explain why the analysis is sound, why the recommendations follow from the data.

Don't. Every defense gives the colleague another surface to attack. Instead: "It sounds like you've been frustrated about this for a while" (CARS: Connect). "There are a couple of ways we could address the concerns you're raising" (Analyze). If a specific factual claim was made — "You didn't even look at the Q3 data" — one correction: "I did include the Q3 data on slide seven" (Respond). Then: "I'd like to hear the specific issues you see — can we set up thirty minutes this week to go through them?" (Set limits by offering a bounded alternative).

What you did not do: defend yourself, explain your reasoning, argue that the presentation was good, respond to the emotional content as if it were intellectual content. What you did: de-escalated, preserved your own dignity, gave the colleague a path forward that doesn't require them to concede anything in public, and — crucially — moved the conversation from a public forum where the HCP's audience is present to a private one where the dynamics are different.

9.3.5 *The Realistic Goal*

For genuine HCPs, conversion to joint inquiry is usually not achievable. Eddy is explicit: some people represent pre-dialogical conditions. The structural preconditions for honest dialogue — the capacity to acknowledge uncertainty, to track the other's reasoning, to update on evidence — are impaired or absent. No technique list repairs a personality structure.

The realistic goals narrow substantially:

Containment means keeping communications limited and factual, using BIFF, not providing material to extend the conflict. The goal is not productive engagement but minimum-damage interaction.

De-escalation means using CARS and Verbal Judo to reduce emotional intensity in real-time encounters — not resolving the underlying dispute, but lowering the temperature enough to reach a functional conclusion to the interaction.

Protection of the audience (Rauch's framing) means that in any public encounter, the people watching are often reachable even when the HCP is not. Addressing your responses partly to the audience — not conspicuously, but with the awareness that they are there — means the encounter can serve an educational function even when it fails conversationally.

Graceful exit means leaving without contempt, without provocation, without providing the HCP with an attack surface for the next round. The exit is not defeat. It is recognition that the preconditions for conversation are absent and that continuing is counterproductive.

9.3.6 *When to Walk Away*

When two or more WEB signals are active AND the BIFF and CARS approaches have not reduced escalation — when the pattern has repeated across multiple encounters regardless of your approach — you are not dealing with a communication problem. You are dealing with a personality structure problem. The decision becomes relational and contextual rather than conversational: how much contact with this person is structurally necessary, and how can you minimize it?

Eddy's counterintuitive prescription: HCPs are not improved by engaging their grievances. Each engagement that provides new material for the conflict extends the cycle. The minimum necessary engagement, maximum brevity, and no emotional content — this is not cold or disrespectful. It is appropriate recognition that the alternative produces escalation that harms everyone involved, including third parties who observe or are drawn into the pattern.

One ethical note deserves explicit statement. Managing rather than fully engaging an HCP is ethically appropriate when full honest engagement reliably produces escalation that harms the conversation's other participants. This is not manipulation. It is recognition that the preconditions for the ethics of honest dialogue — which include a partner capable of reciprocating — are absent. The ethics of dialogue presuppose a dialogue. When only one party is in a dialogue, the ethical obligations are different.

There is a harder version of this ethical question that the review should not duck: what about the HCP's experience? They are, from their own perspective, genuinely wronged. Their emotional responses, while disproportionate to the external observer, are real to them. The decision to manage rather than engage carries a cost — it means declining to treat the HCP as a full conversational partner, which is, in a certain light, a reduction of their moral status in the interaction. This cost is real and should not be disguised. The justification is consequentialist: full engagement reliably produces outcomes that are worse for everyone, including the HCP. The cost of reduced engagement is real but smaller than the cost of the alternative.

9.4 CLOSING: WHAT THIS COMPETENCE ULTIMATELY REQUIRES

This review has covered substantial territory: six dialogue types, three quality dimensions, ten rules for critical discussion, six categories of playbook moves, and three hard cases organized by convertibility. Before closing, it is worth naming what runs underneath all of it — not as a summary, but as an argument about the nature of the competence itself.

9.4.1 *Three Findings That Converge*

Three integrative findings emerged repeatedly across source traditions that rarely cite each other. The convergence is the best evidence that they are tracking something real rather than merely reflecting the assumptions of a single field.

Non-judgmental listening is the universal precondition. Eight independent traditions arrived at this without citing each other: Motivational Interviewing, Street Epistemology, deep canvassing, crisis intervention, restorative justice, the rationality community, Schein's organizational psychology, and Minson's social psychology research. Itzchakov and DeMarree (2022) provide the mechanism: perceived listening quality predicts attitude change through the mediator of psychological safety. The interlocutor becomes willing to reveal their doubts only when the conversation is safe enough to reveal them.

Every move in the playbook presupposes listening as its foundation. Techniques deployed without genuine listening become manipulation — even when they formally resemble the same techniques deployed with it. The interlocutor generally knows the difference. The Socratic question asked with genuine curiosity and the Socratic question asked as a trap are indistinguishable in their surface grammar and completely distinguishable in how they land.

Santoro et al. (2025, *PNAS*) qualify this finding in a way that matters: high-quality listening improves the speaker's perception of the listener but does not, by itself, change attitudes. Listening is the precondition; it is not the active ingredient. The review that told you "just listen better" would be incomplete. Listening creates the psychological safety in which change becomes possible. Something else has to do the work of actually producing it.

Narrative exchange outperforms argument. Broockman and Kalla's deep canvassing research (2016, *Science*; 2020) is the strongest empirical finding in the entire domain this review covers. When the personal narrative component of the protocol was removed while the argumentative content was preserved, the effect disappeared. The argument-only condition had no impact. The mechanism of change was narrative exchange: sharing personal experiences that produced analogic perspective- taking, not propositional engagement.

The practical implication cuts against the intuition that good arguments should win. Stop presenting evidence as your primary move. Start sharing stories — yours and eliciting theirs. People update their attitudes through lived analogy more reliably than through propositional argument. This is not rhetorical strategy. It is the empirically validated mechanism.

Technique rebuttal outperforms topic rebuttal. Schmid and Betsch (2019, *PNAS*) formalized what Street Epistemology practitioners had been discovering empirically: addressing the HOW of belief formation rather than the WHAT of belief content is more effective. The entry point for productive conversation is rarely the contested claim. It is the epistemic process: How did you come to believe this? How confident are you, and why? What would change your mind?

These three findings share a structural feature that is easy to miss because they are usually presented separately. All three involve the same cognitive shift: from third-person to first-person. Listening shifts the speaker from performing a position to examining it. Narrative exchange shifts the conversation from "the evidence says" (third-person, abstract) to "here's what happened to me" (first-person, concrete). Epistemological questioning shifts the focus from "is this claim true?" (third-person evaluation) to "how did I come to believe this?" (first-person examination). In each case, the mechanism of change is not information entering from outside but attention turning inward. The person does not update because you persuaded them. They update because the conversation created conditions in which they could examine their own position from a slight distance — and at that distance, they could see what they couldn't see while standing inside it.

This is, in essence, Bohm's proprioception of thought achieved through conversational structure rather than through meditation or self-discipline. The practitioner's job is not to change the other person's mind. It is to create the conditions under which the other person can see their own mind — its commitments, its uncertainties, its assumptions — well enough to decide for themselves whether any of it should change.

9.4.2 *The Competence Stack*

The competence target that this review serves has five layers. It is worth mapping the review's coverage to them, because the review has addressed them unevenly, and honesty about which layers a document can address is itself a form of the intellectual honesty the document recommends.

Knowledge (Layer 1) — the conceptual map: what dialogue types exist, how to recognize them, what each requires. This review provides this directly, in Part I.

Skill (Layer 2) — the playbook: specific named moves, practiced until fluid. This review provides the names and scripts. The practice is the reader's work. No document can substitute for it, any more than a book about swimming can substitute for swimming.

Judgment (Layer 3) — the capacity to read a live situation and select the right move for the game actually being played. Not the game one wishes were being played, or the game one is most comfortable playing, or the game one played successfully last time with a different person. The diagnostic in Part I is a scaffold for this judgment. But judgment develops through application, error, and reflection — through the slow accumulation of pattern recognition that comes from having been wrong about which game you were in and having noticed it in time, or having noticed it too late and learned from the delay.

Metacognition (Layer 4) — the capacity to notice, in real-time, when one has shifted from truth-seeking to defending. When the scout has become the soldier. When the question one is asking has stopped being genuine and has become a trap. When the frustration with the other person's stubbornness has become the very stubbornness one is complaining about. Bohm's proprioception of thought. The review has named this capacity and identified the moves associated with it. But the monitoring itself — catching the moment when you stopped listening and started performing — is a skill that develops slowly and imperfectly, and no technique list accelerates it faster than honest practice.

Character (Layer 5) — the honest disposition toward truth that makes the techniques authentic. This is where the review must be most careful about what it claims, because this is where the temptation to inspirational language is strongest and the warrant for it is weakest.

The character layer is not a set of beliefs about the importance of truth. It is a set of habits — reflexive behaviors that have been practiced until they are automatic. The habit of saying "I don't know" when you don't know, even when not knowing is uncomfortable. The habit of saying "I was wrong" when you were wrong, even when being wrong threatens your position in the conversation. The habit of noticing when you are defending rather than examining, and pausing, even when the pause costs you a tactical advantage. The habit of treating the other person's experience as real data about the world rather than as an obstacle to be managed.

These habits can be cultivated. They are not personality traits that you either have or don't. They are practices — things you do, imperfectly, with increasing frequency, over time. The review cannot install them. What it can do is make the reader aware that they exist, that they are observable in others, and that their absence is detectable — by the interlocutor, by the audience, and, eventually, by oneself.

9.4.3 *What a Good Conversation Actually Is*

A good conversation, by the standards this review has developed, is not one that ends in agreement. It is one that ends with both parties having learned something they did not know before — about the topic, about each other's reasons, or about their own. The best conversations often end in productive disagreement: a clearer statement of what is actually at stake, a more precise understanding of

where the real disagreement lies, and — sometimes — a mutual recognition that the question is harder than either party thought when they began.

That standard is achievable more often than most people expect. It requires less cleverness than most technique books suggest. And it requires more honesty — with one's interlocutor, and about one's own reasoning — than is comfortable on most days.

The playbook gives the reader the moves. The map gives them the orientation. The hard cases give them a realistic sense of what is and is not achievable in the most difficult terrain. But the character layer — the honest disposition without which the map and the playbook become tactical rather than genuine — this is what the reader must bring to the conversation themselves.

No document can supply it. But a good document can make the reader more likely to notice when it's present, when it's absent, and what it feels like to practice it — imperfectly, as all practice is — in conversations where it matters.

REFERENCES

Aristotle. (2019). *Aristotle's "Art of Rhetoric"* (R. C. Bartlett, Trans.). University of Chicago Press.

Boghossian, P., & Lindsay, J. (2019). *How to Have Impossible Conversations: A Very Practical Guide*. Lifelong Books / Hachette.

Bohm, D. (1996). *On Dialogue*. Routledge. (Ed. Lee Nichol; based on 1990 seminars.)

Broockman, D., & Kalla, J. (2016). Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science*, 352(6282), 220–224.

Cook, J., Lewandowsky, S., & Ecker, U. K. H. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLOS ONE*, 12(5), e0175799.

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., ... & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1, 13–29.

Eddy, B. (2014). *BIFF: Quick Responses to High-Conflict People, Their Hostile Emails, Personal Attacks, and Social Media Meltdowns*. Unhooked Books.

Eddy, B. (2018). *5 Types of People Who Can Ruin Your Life: Identifying and Dealing with Narcissists, Sociopaths, and Other High-Conflict Personalities*. TarcherPerigee.

van Eemeren, F. H., & Grootendorst, R. (1992). *Argumentation, Communication, and Fallacies: A Pragma-Dialectical Perspective*. Lawrence Erlbaum Associates.

Fisher, R., & Ury, W. (2011). *Getting to Yes: Negotiating Agreement Without Giving In* (3rd ed.). Penguin Books. (Original 1981.)

Galef, J. (2021). *The Scout Mindset: Why Some People See Things Clearly and Others Don't*. Portfolio / Penguin.

Grapsas, S., Brummelman, E., Back, M. D., & Denissen, J. J. A. (2019). The "why" of narcissism: A process model of narcissistic status pursuit. *Psychological Bulletin*, 146(2), 150–165.

Haidt, J. (2012). *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. Pantheon Books.

Itzchakov, G., & DeMarree, K. G. (2022). The links between listening, psychological safety, and attitude change. *Frontiers in Psychology*, 13, 930492.

Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.

Kalla, J. L., & Broockman, D. E. (2020). Reducing exclusionary attitudes through interpersonal conversation: Evidence from three field experiments. *American Political Science Review*, 114(2), 410–425.

Lundahl, B., Kunz, C., Brownell, C., Tollefson, D., & Burke, B. (2010). A meta-analysis of Motivational Interviewing: Twenty-five years of empirical studies. *Research on Social Work Practice*, 20(2), 137–160.

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34(2), 57–74.

Miller, W. R., & Rollnick, S. (2013). *Motivational Interviewing: Helping People Change* (3rd ed.). Guilford Press.

Minson, J. A., Hagmann, D., & Luo, L. (2023). Quantifying conversational receptiveness in disagreements: A multi-study investigation. Preprint / Working paper.

Minson, J. A., & Yeomans, M. (2021). Conversational receptiveness: Improving disagreement with a difference in perceived open-mindedness. *Organizational Behavior and Human Decision Processes*, 164, 135–148.

Nguyen, C. T. (2018). Echo chambers and epistemic bubbles. *Episteme*, 17(2), 141–161.

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2), 303–330.

Perelman, C., & Olbrechts-Tyteca, L. (1969). *The New Rhetoric: A Treatise on Argumentation* (J. Wilkinson & P. Weaver, Trans.). University of Notre Dame Press. (Original French: 1958.)

Rauch, J. (2021). *The Constitution of Knowledge: A Defense of Truth*. Brookings Institution Press.

Santoro, E., Broockman, D. E., Kalla, J. L., & Porat, R. (2025). A field experiment testing whether listening improves persuasion. *Proceedings of the National Academy of Sciences*, 122(2).

Schein, E. H. (2013). *Humble Inquiry: The Gentle Art of Asking Instead of Telling*. Berrett-Koehler.

Schmid, P., & Betsch, C. (2019). Effective strategies for rebutting science denialism in public discussions. *Proceedings of the National Academy of Sciences*, 116(31), 15534–15540.

Stone, D., Patton, B., & Heen, S. (2010). *Difficult Conversations: How to Discuss What Matters Most* (rev. ed.). Penguin Books. (Original 1999.)

Storr, W. (2021). *The Status Game: On Social Position and How We Use It*. William Collins.

Tavris, C., & Aronson, E. (2020). *Mistakes Were Made (But Not by Me): Why We Justify Foolish Beliefs, Bad Decisions, and Hurtful Acts* (3rd ed.). Mariner Books. (Original 2007.)

Thompson, G. J., & Jenkins, J. B. (2013). *Verbal Judo: The Gentle Art of Persuasion* (updated ed.). William Morrow. (Original 1993.)

Ury, W. (1991). *Getting Past No: Negotiating in Difficult Situations*. Bantam Books.

Voss, C., & Raz, T. (2016). *Never Split the Difference: Negotiating as if Your Life Depended on It*. HarperBusiness.

Walton, D. (2013). *Methods of Argumentation*. Cambridge University Press.

Walton, D., & Krabbe, E. C. W. (1995). *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. SUNY Press.

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry*, 14(3), 270–277.

Wood, T., & Porter, E. (2019). The elusive backfire effect: Mass attitudes' steadfast factual adherence. *British Journal of Political Science*, 49(4), 1414–1433.