

Structured Challenge, Prediction Failure, and Expectancy: A Mechanistic Account of How Development Practitioners Disrupt Epistemic Rigidity in the 3B Behavior Modification Model

Dr. David M. Robertson, MSL, VL2

GrassFire Industries LLC, Wichita, Kansas, United States

ORCID: <https://orcid.org/0009-0008-0102-1379>

Abstract

Leadership and cognitive-behavior development routinely fail for a reason that is rarely named with precision. Pep talks raise surface motivation without altering the priors that generate behavior. Self-help fails because the mind either protects existing models or cannot see the models it is using. This paper argues that three well-documented phenomena (prediction failure, the nocebo effect, and the Pygmalion effect, with its negative twin, the Golem effect) intersect at a single operating point: expectancy sets the prior against which error is computed, and only a structured challenge can make that error usable. Epistemic Rigidity Theory describes the protective architecture that down-weights disconfirming evidence. The 3B Behavior Modification Model locates the intervention at bias rather than at belief or behavior. Without training in those frameworks, practitioners default to affirmation, information, or unstructured pressure. Those methods do not move the needle because they do not force a specific prediction to fail under constraints the person cannot privately rewrite. The paper specifies the mechanism, distinguishes structured challenge from mere difficulty, and states why development that avoids this collision is training or entertainment rather than correction. The joint model predicts that affirmation plus information will not move bias, and that a designed collision in the named domain, inside a pre-specified predicted-success band, under witness, and repeated once, will. If those two results do not separate, the integration is a re-description and the claim fails.

Keywords: epistemic rigidity, 3B behavior modification model, prediction error, nocebo effect, Pygmalion effect, Golem effect, structured challenge, contrastive inquiry, leadership development, expectancy

Introduction

Development practitioners inherit a market that sells comfort as growth. Workshops affirm identity. Coaches validate perspective. Books supply techniques the reader can adopt without ever discovering which prediction is running the show. Temporary energy is then mistaken for change. When old patterns return under load, the failure is attributed to insufficient motivation or insufficient information. That diagnosis is wrong.

The brain is a prediction engine. It generates expectancies, compares them with incoming evidence, and uses the mismatch as a teaching signal (Schultz, 2016; Friston, 2010). Learning is not the arrival of new content. Learning is the revision of a prior after a prediction fails at sufficient magnitude and precision. If the prior is protected, the same mismatch is treated as noise, attributed to circumstance, or used to justify

the original model. Epistemic Rigidity Theory names that protective architecture as a layered interaction of biases rather than as a single stubborn trait (Robertson, 2024).

Expectancy is the content of the prior. Interpersonal expectancy research shows that one person's prediction about another can alter climate, challenge, airtime, and feedback, and thereby alter performance (Rosenthal & Jacobson, 1968; Eden, 1992). Negative expectancy produces the opposite movement, often called the Golem effect (Babad et al., 1982). Intrapersonal negative expectancy produces nocebo: anticipated harm or incompetence that becomes physiology and behavior (Scott et al., 2008; Colloca & Barsky, 2020). These literatures have been kept in separate rooms. They belong in one mechanism.

This paper advances a single claim. Prediction failure is necessary but not sufficient. Nocebo and Pygmalion set the valence and source of the prior. Structured challenge is the experimental design that forces those priors to collide with observable evidence in a domain, at a dose, under witness, and across repetition. Practitioners trained in Epistemic Rigidity Theory and the 3B Behavior Modification Model can build that design. Practitioners who are not so trained tend to talk at beliefs, soothe emotion, or raise difficulty without closing escape routes. Those moves leave the bias intact.

Prediction Failure as the Teaching Signal

Predictive processing treats perception and action as inference under uncertainty. The system minimizes surprise by updating its generative model or by acting to make the world match the model (Friston, 2010). Prediction error is the residual after that comparison. Dopamine systems encode reward prediction error and support updating when outcomes deviate from what was anticipated (Schultz, 2016). The magnitude of the error matters. Experimental work on belief revision shows that larger prediction errors produce larger updates than the same evidence presented without a prior prediction, and that this relationship holds across partisan and nonpartisan content (Vlasceanu et al., 2021).

That finding is easy to misuse. It does not say that flooding a person with contradiction produces change. It says that a prediction must be made, then broken, and that the size of the break tracks the size of the update when the error is allowed to register. Epistemic rigidity is precisely the condition in which the error is not allowed to register at full precision. In an fMRI pain paradigm with healthy adults, anxiety associated with negative expectancy suppressed prediction-error signaling and thereby helped perpetuate the negative prior (Tsai et al., 2024). The finding is specific to that domain. The transferable point is narrower: threat-weighted negative expectancy can reduce the precision of the error that would otherwise update the model. The person experiences the mismatch and then explains it away. The model survives because the weighting, not the data, is the problem.

In development settings this appears as a familiar sequence. The client hears a better account of their situation. They nod. They leave with language. Under the next real constraint they run the old prediction. The practitioner concludes that they were not ready. A more accurate reading is that no prediction was forced to fail in a form the existing model could not absorb.

Nocebo: The Intrapersonal Negative Prior

The nocebo effect is the production of adverse outcomes by negative expectancy. It is not a reporting bias alone. Placebo and nocebo responses track opposite dopaminergic and endogenous opioid responses in networks implicated in reward and motivation. Nocebo is associated with deactivation of those systems (Scott et al., 2008). Reviews of clinical and experimental work show that verbal suggestion is often sufficient to generate nocebo, and that situationally induced expectancy and state anxiety, rather than trait anxiety, predict the size of the effect (Colloca & Barsky, 2020; Rooney et al., 2023).

Translated out of the clinic, the same structure appears as an analogue of nocebo: the client's high-precision prediction that development will not work for them, that they cannot perform under this constraint, or that any apparent gain will collapse. Clinical nocebo is defined by symptom and physiological endpoints. The leadership analogue shares the computational form (a negative prior that shapes sampling and outcome) without yet sharing those endpoints. Subsequent uses of the term in this paper refer to that analogue. The prediction is frequently inherited from prior failed programs, from labels, and from social models. Once installed, it changes what counts as evidence. Ambiguous outcomes are coded as confirmation. Early gains are coded as luck. The practitioner's optimism is coded as naivete. This is not a motivation deficit. It is a generative model that predicts harm or incapacity and then samples the world accordingly.

Expectancy effects can become self-reinforcing: once a prior is in place, subsequent experience is interpreted through it, which maintains the effect unless the expectancy itself is disrupted (Jepma et al., 2018). That is useful and incomplete. A single violation can be reclassified. Without structure, the nocebo prior reasserts itself by relocating the failure: wrong timing, wrong audience, wrong metric. The mind protects itself. That protection is the point of Epistemic Rigidity Theory (Robertson, 2024).

Pygmalion and Golem: The Interpersonal Prior

The Pygmalion effect is interpersonal expectancy operating as a self-fulfilling prophecy. In the original classroom work, teachers were led to expect growth from randomly designated students, and those students later showed measured gains (Rosenthal & Jacobson, 1968). Differential treatment was inferred from the outcome pattern rather than recorded as a manipulated mediator, which is part of why the study was immediately contested (Thorndike, 1968). Organizational extensions treat leader expectancy as a cause of follower performance through changes in leader behavior (Eden, 1992). The negative form, the Golem effect, describes performance decline under low expectancy (Babad et al., 1982). Rosenthal's later four-factor account specifies the mediating channels: climate (warmth and regard), input (how much and how hard the material), output opportunity (how often the person is invited to act and how long the other waits), and feedback (specificity and richness) (Rosenthal, 1994). The four channels are not equally weighted. Climate and input carry the larger mediating effects in the meta-analytic evidence, with output opportunity and feedback contributing less, which tells a practitioner where to spend attention first.

Honesty about magnitude is required. The original intelligence claims were contested, replications were mixed, and later analyses argue that much of the observed expectancy–outcome correlation is accurate perception rather than causal transformation (Thorndike, 1968; Jussim & Harber, 2005). Field attempts to train a Pygmalion leadership style have often produced small effects (Eden et al., 2000). That does not erase

the mechanism. It constrains the slogan. A practitioner who merely decides to believe in the client, without changing climate, input, opportunity, and feedback, is not running Pygmalion. They are running a pep talk.

The interpersonal prior still matters because it changes the evidence the client will use to test their own nocebo. If the practitioner expects little, they withhold stretch, interrupt sooner, and give thinner feedback. The client's prediction of incapacity is then confirmed by the situation the practitioner built. Golem and nocebo then lock. If the practitioner expects more but never constructs a collision the client cannot privately rewrite, Pygmalion remains a private attitude. The client never has to update.

The Intersection

The three components meet at one computational moment. The client carries a prior, often negatively valenced and high in precision. The practitioner carries a prior about the client's capacity. Behavior in a defined task produces an outcome. That outcome is a prediction error relative to one or both priors. Whether the error revises anything depends on weighting.

If the client's nocebo prior is more precise than the incoming evidence, the error is discounted. If the practitioner's Golem prior shapes the task so that the client never receives a clean disconfirmation, there is no error worth weighting. If both priors are positive but the task is unstructured, the client can grant the success to special conditions and keep the old identity. Only a designed collision produces a signal that is both large enough and constrained enough to force reweighting.

This is why pairwise treatments miss the applied problem. Prediction-error research shows that surprise can update beliefs when people first commit to a prediction (Vlasceanu et al., 2021). Nocebo research shows that negative expectancy can suppress or distort that updating and can alter physiology (Scott et al., 2008; Tsai et al., 2024). Pygmalion research shows that another person's expectancy changes the data stream. Development requires all three to be handled at once: the client's prior, the practitioner's prior, and the designed mismatch.

The joint model is not a translation of the three literatures into one vocabulary. It makes predictions that no component makes alone. Prediction-error research predicts that a larger mismatch produces a larger update when the error is allowed to register. It does not predict that the same mismatch will fail to update when the task sits in an adjacent domain, when witness is absent, or when a single hit can be reclassified as luck. Nocebo research predicts that a high-precision negative prior will distort sampling and can suppress the error signal. It does not specify the interpersonal channels that keep supplying confirming evidence. Pygmalion research predicts that another person's expectancy changes climate, input, opportunity, and feedback. It does not predict that high expectancy without a closed collision will leave the client's prior intact. The joint model does. Affirmation plus information, with no named prediction, no pre-specified predicted-success band, no witness, and no second trial, should not move bias. A designed collision in the named domain, inside that band, under witness, and repeated once, should. If those two results do not separate, the integration is a re-description and the claim fails.

Why Pep Talks and Self-Help Do Not Touch the Mechanism

Pep talks operate on emotion and on stated belief. They tell the person they are capable. However, they do not make a specific prediction fail. The 3B Behavior Modification Model holds that emotion drives bias, bias drives belief, belief drives behavior, and behavior drives outcomes (Robertson, 2025a). Targeting belief while leaving bias untouched produces compliance language without model revision. The emotional lift of affirmation can even increase precision on the existing identity: the person feels supported in remaining who they already think they are.

Self-help fails for two additional reasons. First, the person cannot see the prior they are using. Epistemic rigidity includes Dunning–Kruger dynamics, Einstellung defaults, anchoring, and confirmation loops that hide the model from its owner (Robertson, 2024). A person cannot correct a prediction they do not know they are making. Second, when a glimpse of the model appears, protection starts. Cognitive dissonance, motivated reasoning, and threat-related anxiety reduce the precision of the contradictory evidence. Unsupervised reading supplies information. It does not supply a constrained test with a witness.

Unstructured pressure is not a solution. Harder tasks without closed escape routes produce either flooding or theater. Flooding raises threat precision and suppresses the error signal. Theater produces a performance the person can later disown. In both cases the bias remains. At the same time, it is critical to understand that difficulty is not structure.

Structured Challenge as the Corrective

Structured challenge is the independent variable. Prediction failure is the dependent variable the structure is built to produce. The structure has four necessary constraints.

Domain specificity requires that the task hit the exact prior doing the damage. A client who predicts they cannot hold a hard conversation is not corrected by succeeding at a spreadsheet. Adjacent competence is how rigidity survives. The challenge must be in the channel where the nocebo lives. Hence, the exact nocebo must be identified.

Dose requires a mismatch large enough that assimilation is costly and small enough that anxiety does not silence the signal. Too little challenge leaves predictions intact. Excessive, chronic, or uncontrollable adversity damages the systems required for updating rather than recruiting them (McEwen & Morrison, 2013; Maier & Seligman, 2016). The productive window is moderate, time-limited, and recoverable. That window is what a trained practitioner is for. A trained practitioner must selectively choose the model that best fits the client's needs while ensuring the selected method targets the weakest points of their Epistemic Rigidity. For testing, dose needs a candidate operationalization rather than unaided judgment alone. Dose is a property of the task as designed, not of the outcome the task produces, so it has to be fixed before the trial runs. The workable operationalization is a target band: the practitioner calibrates the task so that the client's pre-trial predicted-success rating on the named criterion falls inside a stated range, and that range is the dose. A post-trial difficulty rating anchored to the stated criterion rather than to global discomfort can then be recorded as a manipulation check. The size of the miss and the first-trial failure rate are outcomes and belong on the dependent side of the design. Without a pre-specified band, a comparison against affirmation-plus-information cannot isolate dose.

Witness requires evidence the person cannot privately re-narrate. A prediction stated in advance, a criterion stated in advance, and an observer who saw the outcome close the usual exits. Pygmalion is not a vibe in this design. It is the practitioner's public expectancy that the person can meet the criterion under the stated constraint. That expectancy is itself a prediction that will be confirmed or broken in the same trial.

Iteration requires that the same class of prediction be tested again after the first miss or first hit. One exception can be absorbed as luck. A second trial under the same rules changes the cost of keeping the old precision. Belief update tracks repeated weighting, not a single anecdote.

When those constraints are in place, the 3B sequence can move at the correct layer. The failed prediction produces emotion. That emotion is not soothed away. It is used as the signal that the bias (the weighting rule) is now expensive. Discomfort here is a benefit. Once weighting shifts, belief follows because the old belief no longer pays for itself. Behavior changes without exhortation because the person is no longer predicting the old outcome with high precision. Outcomes then supply further evidence. Contrastive Inquiry is the questioning method that keeps the contrast in view so the person cannot slide back to a single frame (Robertson, 2025b). The method generates a contradictory hypothesis in order to examine it. What it does not do is hand the client a forced choice between two pre-selected framings and call the choice an update. The contrast must be produced and tested, not administered as a binary menu.

This account is consistent with the broader claim that development must be uncomfortable. Neural remodeling follows a prolonged mismatch between current capacity and environmental demand (Lövdén et al., 2010), and that demand must exceed capacity while remaining survivable (McEwen & Morrison, 2013). Comfortable programs conserve wiring. Programs that break people produce degradation. Only calibrated collision produces correction. The Adversity Nexus describes the social and organizational cycle in which the removal of productive adversity precedes stagnation (Robertson, 2023). That cycle is the context in which unstructured comfort becomes expensive. It is not itself a neural model.

Implications for Practitioners Trained in 3B and Epistemic Rigidity

Training in these frameworks is not ornamental. It changes what the practitioner watches and what they refuse to do. This path offers an alternative.

The practitioner watches for the client's predicted failure, not only for the client's stated goal. They listen for nocebo phrasing masked as realism. They notice their own Golem leakage in wait time, task selection, and the richness of feedback. They do not treat a warm session as progress.

The practitioner refuses three defaults. They refuse affirmation as the primary intervention. They refuse information as the primary intervention. They refuse difficulty without a closed criterion. Each default leaves the prior untested.

The practitioner builds micro-trials. A trial names the prediction, names the constraint, names the success criterion, runs the behavior, and records the outcome in front of both parties. That is an experiment, not a conversation about identity. If the prediction fails in the client's favor, the practitioner does not celebrate personality. They ask what the old model must now reduce in precision. If the prediction is confirmed, they

do not moralize. They adjust dose or domain and run the next trial. Failure of a trial is data. Failure to structure a trial is malpractice relative to the mechanism.

Without this understanding, failure is likely for a mechanical reason. The helper's tools never reach bias. Pep talks never begin to touch the weighting rule. Self-help never sees the rule or else defends it. The mind protects itself because it is designed to, or it simply does not know what it does not know. Those are the two doors through which expensive programs exit without residue.

Limitations and Scope

This paper is a theoretical integration, not a new randomized trial of the triad as a packaged protocol. Each component has an independent evidence base. The joint operating model is an argument from mechanism and should be tested as such: pre-registered prediction, structured task, an operationalized dose (a pre-specified predicted-success band, with criterion-anchored difficulty as a manipulation check), and repeated trials against a comparison condition of affirmation-plus-information.

The primary outcome has to sit at the layer the 3B model targets. Bias is the weighting rule. Belief is the defended report. A global expectancy or identity rating is a belief item, and treating it as the test of change would repeat the failure mode this paper assigns to pep talks. The primary outcome is a shift in predicted-success rating on a fresh unrehearsed task in the same domain, together with performance on that task. Practitioner and client ratings of global expectancy or identity remain secondary.

The first study is a package test. It can show that the four-constraint design outperforms affirmation-plus-information. It cannot show that each constraint is necessary, which is what the claim requires. The follow-on is a dismantling design: the same protocol with one constraint dropped per arm, so that an arm without witness and an arm without a second trial can be compared with the full bundle. If dropping a constraint does not degrade the primary outcome, that constraint is not doing the work assigned to it.

Pygmalion should not be over-claimed. Effect sizes in applied leadership training are often small, and accurate perception explains a substantial share of expectancy–performance correlations (Jussim & Harber, 2005; Eden et al., 2000). The claim here is narrower. Interpersonal expectancy changes the evidence stream through identifiable channels. Those channels can be used or wasted.

Nocebo evidence is strongest in symptom and performance domains with clear physiological or behavioral endpoints. Extension to leadership identity is analogical and should be marked as such. However, the analogy is tight enough to guide practice and still requires direct tests in development samples.

Structured challenge can be done badly. Poor calibration reproduces learned helplessness rather than update (Maier & Seligman, 2016). The ethics of the method require consent to be pushed, permission to be explored, competence to read threat versus usable error, and a recovery interval. Challenge without those conditions is not this paper's recommendation.

Conclusion

Prediction failure, nocebo, and Pygmalion are not three metaphors for trying harder. They are one mechanism viewed from the error signal, the intrapersonal prior, and the interpersonal prior. Epistemic rigidity is what happens when that mechanism is allowed to protect the old weighting. The 3B model states where to intervene: at bias, after emotion has been recruited by a real miss, before belief is argued with.

Structured challenge is how a practitioner turns the mechanism on purpose. Domain, dose, witness, and iteration convert surprise into a usable error. Absent that design, development remains a pep rally or a library. The person leaves with feeling or with facts. The prior that generates their outcomes does not move.

Practitioners who understand this intersection can, and should, stop confusing comfort with progress. Practitioners who do not will likely keep failing for reasons that look like client resistance, when they are, in fact, an untested model defending itself.

References

- Babad, E. Y., Inbar, J., & Rosenthal, R. (1982). Pygmalion, Galatea, and the Golem: Investigations of biased and unbiased teachers. *Journal of Educational Psychology*, 74(4), 459–474. <https://doi.org/10.1037/0022-0663.74.4.459>
- Colloca, L., & Barsky, A. J. (2020). Placebo and nocebo effects. *The New England Journal of Medicine*, 382(6), 554–561. <https://doi.org/10.1056/NEJMr1907805>
- Eden, D. (1992). Leadership and expectations: Pygmalion effects and other self-fulfilling prophecies in organizations. *The Leadership Quarterly*, 3(4), 271–305. [https://doi.org/10.1016/1048-9843\(92\)90018-B](https://doi.org/10.1016/1048-9843(92)90018-B)
- Eden, D., Geller, D., Gewirtz, A., Gordon-Terner, R., Inbar, I., Liberman, M., Pass, Y., Salomon-Segev, I., & Shalit, M. (2000). Implanting Pygmalion leadership style through workshop training: Seven field experiments. *The Leadership Quarterly*, 11(2), 171–210. [https://doi.org/10.1016/S1048-9843\(00\)00042-4](https://doi.org/10.1016/S1048-9843(00)00042-4)
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138. <https://doi.org/10.1038/nrn2787>
- Jepma, M., Koban, L., van Doorn, J., Jones, M., & Wager, T. D. (2018). Behavioural and neural evidence for self-reinforcing expectancy effects on pain. *Nature Human Behaviour*, 2(11), 838–855. <https://doi.org/10.1038/s41562-018-0455-8>
- Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9(2), 131–155. https://doi.org/10.1207/s15327957pspr0902_3
- Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin*, 136(4), 659–676. <https://doi.org/10.1037/a0020080>
- Maier, S. F., & Seligman, M. E. P. (2016). Learned helplessness at fifty: Insights from neuroscience. *Psychological Review*, 123(4), 349–367. <https://doi.org/10.1037/rev0000033>

- McEwen, B. S., & Morrison, J. H. (2013). The brain on stress: Vulnerability and plasticity of the prefrontal cortex over the life course. *Neuron*, 79(1), 16–29. <https://doi.org/10.1016/j.neuron.2013.06.028>
- Robertson, D. M. (2023, August 28). The Adversity Nexus Theory. *The Journal of Leaderology and Applied Leadership*. <https://jala.nlainfo.org/the-adversity-nexus-theory/>
- Robertson, D. M. (2024, June 26). *Epistemic rigidity: A theoretical framework for understanding cognitive barriers to knowledge advancement* [Preprint]. SSRN. <https://doi.org/10.2139/ssrn.5875142>
- Robertson, D. M. (2025a, February 18). *The 3B behavior modification model: A framework for understanding and reshaping bias-driven behavior* [Preprint]. SSRN. <https://doi.org/10.2139/ssrn.5875502>
- Robertson, D. M. (2025b, February 3). Unlocking understanding with the Contrastive Inquiry Method. *The Journal of Leaderology and Applied Leadership*. <https://jala.nlainfo.org/unlocking-understanding-with-the-contrastive-inquiry-method/>
- Rooney, T., Sharpe, L., Todd, J., Richmond, B., & Colagiuri, B. (2023). The relationship between expectancy, anxiety, and the nocebo effect: A systematic review and meta-analysis with recommendations for future research. *Health Psychology Review*, 17(4), 550–577. <https://doi.org/10.1080/17437199.2022.2125894>
- Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science*, 3(6), 176–179. <https://doi.org/10.1111/1467-8721.ep10770698>
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. Holt, Rinehart and Winston.
- Schultz, W. (2016). Dopamine reward prediction error coding. *Dialogues in Clinical Neuroscience*, 18(1), 23–32. <https://doi.org/10.31887/DCNS.2016.18.1/wschultz>
- Scott, D. J., Stohler, C. S., Egnatuk, C. M., Wang, H., Koeppe, R. A., & Zubieta, J.-K. (2008). Placebo and nocebo effects are defined by opposite opioid and dopaminergic responses. *Archives of General Psychiatry*, 65(2), 220–231. <https://doi.org/10.1001/archgenpsychiatry.2007.34>
- Thorndike, R. L. (1968). Review of the book *Pygmalion in the classroom*, by R. Rosenthal & L. Jacobson. *American Educational Research Journal*, 5(4), 708–711. <https://doi.org/10.3102/00028312005004708>
- Tsai, H.-Y., Lapanan, K., Lin, Y.-H., Huang, C.-W., Lin, W.-W., Lin, M.-M., Lu, Z.-L., Lin, F.-S., & Tseng, M.-T. (2024). Integration of prior expectations and suppression of prediction errors during expectancy-induced pain modulation: The influence of anxiety and pleasantness. *The Journal of Neuroscience*, 44(17), e1627232024. <https://doi.org/10.1523/JNEUROSCI.1627-23.2024>
- Vlasceanu, M., Morais, M. J., & Coman, A. (2021). The effect of prediction error on belief update across the political spectrum. *Psychological Science*, 32(6), 916–933. <https://doi.org/10.1177/0956797621995208>