INTELLIGENCE · ANALYSIS AND TRADECRAFT

Cognitive Bias Mitigation Training and Its Measurable Effects on Analytic Knowledge

Intelligence and business organisations spend considerable effort training analysts to recognise and avoid cognitive biases. This article reviews what research shows about which kinds of training work, how large and lasting their effects are, why awareness alone is not enough and how to measure whether training improves real judgements.

ModQu research article · 2,677 words · 24 references (APA 7th edition)

Since the 1970s, research on judgement has documented a long list of **cognitive biases**: systematic ways in which human reasoning departs from what logic or probability would require. Tversky and Kahneman (1974) showed how simple mental shortcuts, such as judging probability by how easily examples come to mind, lead to predictable errors. Heuer (1999) brought this research to intelligence analysis, arguing that analysts are as vulnerable to bias as anyone and that their errors can have serious consequences.

Organisations have responded with training. Intelligence agencies, militaries, banks, consulting firms and research companies run courses on cognitive bias, and many analysts can name a dozen biases. The question this article addresses is whether such training works: does it measurably improve the quality of analysts' judgements, and if so, which kinds of training, for how long and under what conditions?

The biases that matter most for analysis

Hundreds of biases have been named, but a smaller set is particularly relevant to intelligence and research work.

- **Confirmation bias**: the tendency to seek, interpret and remember information in ways that support existing beliefs. Nickerson (1998) described it as perhaps the most pervasive problem in human reasoning.
- **Anchoring**: the tendency for an initial number or idea to pull subsequent estimates towards it, even when the anchor is irrelevant (Tversky & Kahneman, 1974).
- **Overconfidence**: Moore and Healy (2008) distinguish overestimation of one's own performance, overplacement relative to others and **overprecision**, excessive certainty about the accuracy of one's beliefs. Overprecision is especially common in forecasts and estimates.
- **Availability**: judging likelihood by how easily examples come to mind.
- **Representativeness and base-rate neglect**: judging probability by similarity to a stereotype and ignoring how common outcomes are in general.
- **Hindsight bias**: once an outcome is known, believing it was more predictable than it was.

An additional bias makes all the others harder to fix. Pronin et al. (2002) found that people readily see bias in others but believe themselves to be less biased than average, a pattern they called the **bias blind spot**. Analysts who learn about biases may therefore conclude that the lessons apply mainly to their colleagues.

Why awareness is not enough

The simplest form of training tells people about biases and urges them to avoid them. Research suggests this has little effect. Fischhoff (1982) reviewed early attempts to reduce hindsight bias and overconfidence and found that warnings alone, even explicit ones, produced little improvement. More intensive interventions, such as personalised feedback and extended training, were needed. Heuer (1999) reached a similar conclusion for analysts: knowing about a bias does not make one immune to it, because biases operate automatically and largely outside awareness.

Arkes (1991) explained why different biases respond to different remedies. Some errors arise because people use a poor strategy when a better one is available; these respond to incentives and instruction. Others arise from automatic associations in memory, such as anchoring and confirmation bias; incentives do not help much, because people try harder using the same flawed process. These need techniques that change what information comes to mind.

Two broad approaches

Larrick (2004) and Soll et al. (2015) review the debiasing literature and distinguish two broad approaches.

The first is to **change the person**: to teach skills, rules or habits that analysts carry with them. Examples include training in statistical reasoning, in considering alternatives and in calibration.

The second is to **change the environment**: to redesign procedures, tools and incentives so that good judgement is easier and bias less likely, regardless of what the analyst remembers from training. Examples include checklists, requirements to record probability estimates, independent assessments before group discussion and decision aids.

The two approaches are complementary. Training helps analysts understand why procedures matter; procedures ensure that good practice continues when training fades.

What the evidence says about ways to reduce bias

Some approaches have strong support; others are popular but weakly tested

Approach	What it involves	Evidence
Warning about bias	Tell people biases exist	Weak: little effect on its own
Consider the opposite	Ask why the judgement could be wrong	Good, in many experiments
Interactive training	Games and exercises with feedback	Good; effects last for weeks
Probability training	Base rates, calibration, updating	Good, in forecasting tournaments
Structured techniques	Formal methods such as ACH	Mixed; few controlled tests
Changing the process	Checklists, independent views, scoring	Good, and does not rely on memory

Compiled from Fischhoff (1982), Lord et al. (1984), Morewedge et al. (2015), Mellers et al. (2014), Dhami et al. (2019) and Soll et al. (2015); table by ModQu

ModQu.com

Figure 1. What the evidence says about ways to reduce bias. Warnings alone achieve little; techniques that change what analysts actually do have stronger support. Compiled from Fischhoff (1982), Lord et al. (1984), Morewedge et al. (2015), Mellers et al. (2014), Dhami et al. (2019) and Soll et al. (2015); table by ModQu.

Techniques with good evidence

Considering the opposite

One of the best-supported debiasing techniques is simple: ask people to consider why their judgement might be wrong. Lord et al. (1984) found that instructing participants to consider the opposite reduced biased assimilation of evidence more effectively than instructions to be fair and unbiased. Koriat et al. (1980) found that asking people to list reasons why their answer might be wrong reduced overconfidence. Mussweiler et al. (2000) showed that considering the opposite reduced anchoring even among car experts estimating the value of a vehicle.

The technique works because it changes which information comes to mind. Instead of searching memory only for reasons that support a view, the analyst is prompted to search for reasons against it. It can be built into analytic procedures at little cost: a standard question in every review, a section in every report on what would change the judgement, or a formal exercise such as a premortem.

Training in statistical reasoning

Fong et al. (1986) taught participants statistical principles, such as the law of large numbers, and found that the training improved reasoning about everyday problems, including problems in areas quite different from those used in the training. This was an important finding, since a common concern about training is that it does not transfer beyond the examples taught.

Interactive training with feedback

Morewedge et al. (2015) tested two forms of debiasing training: an instructional video and an interactive computer game in which participants made decisions, received personalised feedback on their biases and were taught strategies to reduce them. Both reduced biases including confirmation bias, anchoring and the bias blind spot, but the game produced larger effects. The improvements were still present when participants were retested two to three months later. This was one of the first demonstrations that a single training session could produce lasting reductions in bias.

Sellier et al. (2019) then tested whether the game's effects would carry over to a real decision in a different setting. Business students who had played the game were later given a case study in class, presented as an ordinary course exercise, in which confirmation bias would lead to a poor recommendation. Students who had been trained were less likely to choose the biased option than those who had not, suggesting that the training changed real behaviour, not only test scores.

Training in probabilistic forecasting

The largest field evidence comes from forecasting tournaments sponsored by the American intelligence research agency. Mellers et al. (2014) found that about an hour of training in probabilistic reasoning, covering base rates, the use of comparison classes, averaging different estimates and avoiding common biases, improved forecasting accuracy measurably over a full year of questions about world events. Chang et al. (2016) analysed several years of tournaments and confirmed that training improved accuracy, with benefits that persisted across the forecasting season.

These results are important for intelligence because the outcomes were real events, not laboratory problems, and accuracy was measured with proper scoring rules. They show that training can improve the quality of judgements on the kinds of questions analysts actually face.

Techniques with weaker evidence

Structured analytic techniques

Intelligence agencies have invested heavily in **structured analytic techniques**, such as analysis of competing hypotheses, key assumptions checks and devil's advocacy. These techniques are intended partly to reduce bias. However, controlled tests are few. Coulthart (2017) reviewed the evidence for twelve core techniques and found that most had not been rigorously tested, and that results for those that had were mixed.

Analysis of competing hypotheses has received the most attention. Dhami et al. (2019) found that intelligence analysts trained in the technique did not produce more accurate judgements than untrained analysts, and that many did not follow its steps closely. Whitesmith (2020), in a series of experiments, found that the technique did not reliably reduce confirmation bias or the tendency to give extra weight to information received first or last.

These results do not mean that structured techniques are useless. They may improve transparency, make reasoning easier to review and help teams discuss disagreements. But their effect on accuracy and bias should not be assumed.

Correcting judgements after the fact

Another approach accepts that individual judgements will be biased and corrects them statistically. Mandel et al. (2018) found that simple mathematical adjustments to analysts' probability judgements, such as making a set of judgements logically consistent with one another and averaging across analysts, improved accuracy more than some training interventions did. This approach belongs to the "change the environment" family: it improves results without requiring analysts to change how they think.

Why some training fails

Lilienfeld et al. (2009) argued that debiasing research had been surprisingly neglected relative to research documenting biases, and identified several obstacles to effective training:

- **The bias blind spot.** Trainees may accept that biases exist but believe the lessons apply mainly to others (Pronin et al., 2002).
- **Lack of transfer.** Skills learned on textbook examples may not be applied to real problems unless training uses realistic cases.
- **Fading.** Without practice, the effects of training decline over time.
- **Lack of feedback.** Analysts often never learn whether their judgements were right, so they cannot tell whether the training helped.

Kahneman et al. (2021) add a further point: much error in professional judgement is **noise**, random variability between judges or between occasions, rather than bias in a consistent direction. Training aimed at specific biases may do little about noise, which is better reduced by structured procedures, independent judgements and aggregation.

Debiasing teams, not only individuals

Most analysis is produced by teams, and team processes can either reduce or amplify individual biases. When members discuss a question before forming their own views, the first opinion expressed often anchors the rest, and confident speakers carry more weight than their accuracy justifies. Kahneman et al. (2021) recommend that each member record an independent judgement before discussion and that the group then examine the reasons behind any differences. Averaging independent estimates reduces noise, and the discussion of disagreements brings out evidence that individuals had overlooked.

Training can support these processes. Teams that have practised considering the opposite together are more likely to use it in real discussions, and teams that review their past forecasts together learn from one another's errors. Mellers et al. (2014) found that forecasters working in teams, sharing information and challenging each other's reasoning, were more accurate than those working alone, and that training and teaming had separate, additive benefits.

The role of leaders and incentives

Training is unlikely to change behaviour if the organisation rewards the opposite. Analysts who are praised for confident, early calls and criticised for expressing uncertainty will learn to sound certain, whatever their training taught. Arkes (1991) noted that incentives alone do not cure association-based biases, but the reverse is also true: training alone does not overcome incentives that encourage overconfidence.

Leaders play a central part. When managers ask "what would change your mind?" and "how sure are you, in numbers?", and when they treat a well-reasoned forecast that turns out wrong differently from a careless one, they create conditions in which debiasing training can take hold. Soll et al. (2015) emphasise that the most durable improvements usually come from changing the decision environment, and leaders control most of that environment.

Measuring the effects of training

Many organisations evaluate bias training only by asking participants whether they found it useful. Kirkpatrick and Kirkpatrick (2006) proposed four levels of training evaluation, which provide a useful framework.

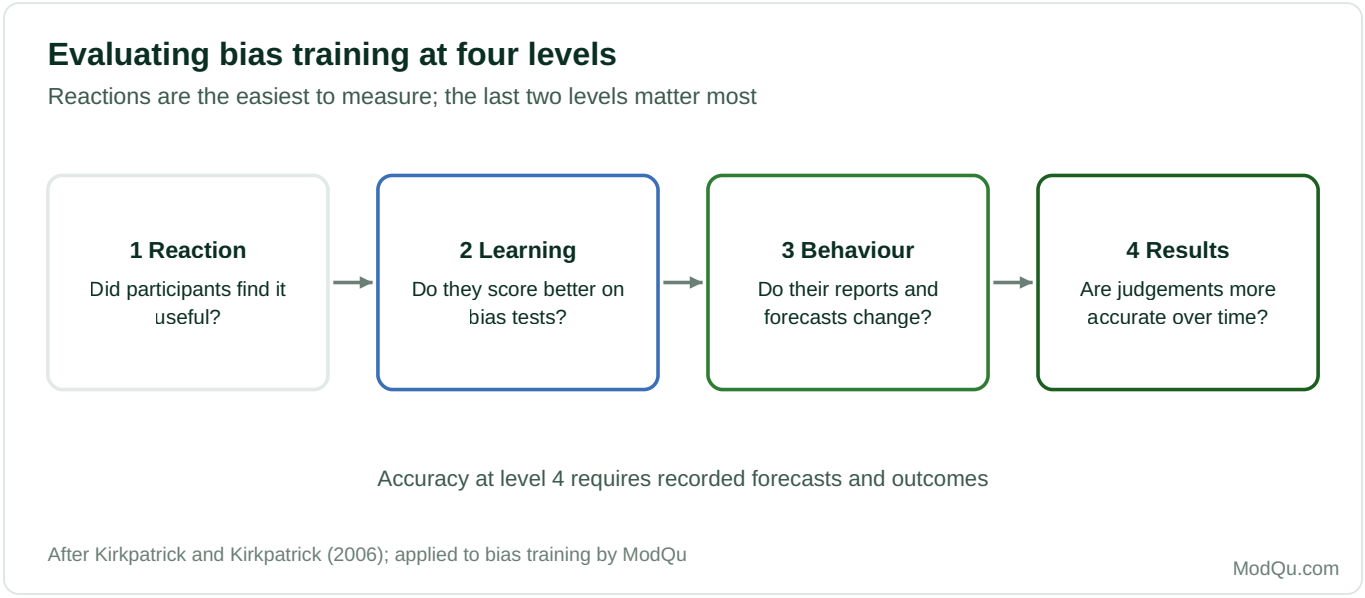


Figure 2. Evaluating bias training at four levels. Measuring real accuracy requires recording forecasts and comparing them with outcomes. After Kirkpatrick and Kirkpatrick (2006); applied to bias training by ModQu.

Reaction is easy to measure but says little about effectiveness. **Learning** can be measured with tests of susceptibility to biases before and after training, as in Morewedge et al. (2015). **Behaviour** requires looking at analysts' actual work: do their reports consider alternatives, state assumptions and express calibrated probabilities? **Results** require measuring the accuracy of judgements over time, which in turn requires analysts to record forecasts in a form that can be scored, as in the forecasting tournaments described by Mellers et al. (2014).

An organisation that wants to know whether its training works should therefore build in measurement from the start: a short bias test before and after training, a review of a sample of reports, and a register of forecasts that can be scored when outcomes are known. Comparing trained and untrained groups, or comparing results before and after training, gives evidence that can guide future investment.

Bias in research and survey analysis

The same biases affect the analysis of survey and market research data. Analysts who expect a product to succeed may focus on the subgroups where it scores well and explain away those where it does not. A first figure in a report, such as last year's market share, can anchor estimates for the coming year. Results that confirm a client's or manager's view may be accepted with less checking than those that contradict it. Practical safeguards include writing down the hypotheses and the planned analysis before the data arrive, asking a colleague to review the results without knowing what was expected, and reporting findings that cut against the main conclusion as prominently as those that support it.

Designing an effective programme

Bringing these findings together, an effective bias mitigation programme for analysts would include the following elements:

1. **Short, interactive training with feedback**, using realistic cases from the analysts' own field rather than abstract puzzles.
2. **Training in probabilistic reasoning**: base rates, calibration, updating and the use of numerical probabilities.
3. **Habitual use of considering the opposite**, built into templates and review procedures.
4. **Process changes** that do not depend on memory: independent judgements before group discussion, checklists for key assumptions and alternatives, and a requirement to record probabilities.
5. **Feedback on accuracy**, through a forecast register that is scored and discussed.
6. **Refresher sessions**, since the effects of training fade without practice.
7. **Measurement at all four levels**, so the programme can be improved.

Common mistakes

- **Teaching lists of biases.** Naming biases is easy; changing behaviour is hard. Training should focus on techniques, not vocabulary.
- **Assuming techniques work.** Structured methods should be tested, not simply adopted (Coulthart, 2017).
- **Ignoring the environment.** Training cannot overcome incentives that reward confidence over accuracy.
- **Measuring only satisfaction.** Positive reactions do not show that judgement has improved.

- **Treating bias training as a one-off event.** Without practice and feedback, gains fade.

An illustration

A Canadian market research company notices that its analysts' sales forecasts for clients are consistently too optimistic and too narrow. It introduces a three-part programme. First, analysts complete a two-hour interactive workshop using past forecasts from the company's own archive, receiving feedback on how their estimates compare with actual results. Second, the forecast template is changed: every forecast must include an 80 per cent range, a statement of the main assumptions and a short note on why the forecast might be too high or too low. Third, all forecasts are entered in a register and scored when results are known.

After a year, the company compares forecasts made before and after the programme. The ranges are wider and capture the actual result more often, and average error has fallen. Analysts report that the "why might this be wrong" note is the most useful single change, which is consistent with the research on considering the opposite.

Conclusion

Training analysts to recognise biases is not enough on its own: awareness has little effect, and people tend to see bias in others more than in themselves. But research shows that some training does work. Considering the opposite, interactive training with personalised feedback and training in probabilistic reasoning all have good evidence behind them, and the effects can last for months and carry over to real decisions. Structured analytic techniques have weaker support than often assumed. The most reliable improvements come from combining training with changes to procedures, tools and feedback, and from measuring the results in terms of real accuracy rather than satisfaction.

References

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin, 110*(3), 486–498. https://doi.org/10.1037/0033-2909.110.3.486

Chang, W., Chen, E., Mellers, B., & Tetlock, P. (2016). Developing expert political judgment: The impact of training and practice on judgmental accuracy in geopolitical forecasting tournaments. *Judgment and Decision Making, 11*(5), 509–526.

Coulthart, S. J. (2017). An evidence-based evaluation of 12 core structured analytic techniques. *International Journal of Intelligence and CounterIntelligence, 30*(2), 368–391. https://doi.org/10.1080/08850607.2016.1230706

Dhami, M. K., Belton, I. K., & Mandel, D. R. (2019). The "analysis of competing hypotheses" in intelligence analysis. *Applied Cognitive Psychology, 33*(6), 1080–1090. https://doi.org/10.1002/acp.3550

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). Cambridge University Press.

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology, 18*(3), 253–292. https://doi.org/10.1016/0010-0285(86)90001-0

Heuer, R. J., Jr. (1999). *Psychology of intelligence analysis*. Center for the Study of Intelligence, Central Intelligence Agency.

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A flaw in human judgment*. Little, Brown Spark.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels* (3rd ed.). Berrett-Koehler.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 107–118. https://doi.org/10.1037/0278-7393.6.2.107

Larrick, R. P. (2004). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 316–337). Blackwell.

Lilienfeld, S. O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on Psychological Science, 4*(4), 390–398. https://doi.org/10.1111/j.1745-6924.2009.01144.x

Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology, 47*(6), 1231–1243. https://doi.org/10.1037/0022-3514.47.6.1231

Mandel, D. R., Karvetski, C. W., & Dhami, M. K. (2018). Boosting intelligence analysts' judgment accuracy: What works, what fails? *Judgment and Decision Making, 13*(6), 607–621.

Mellers, B., Ungar, L., Baron, J., Ramos, J., Gurcay, B., Fincher, K., Scott, S. E., Moore, D., Atanasov, P., Swift, S. A., Murray, T., Stone, E., & Tetlock, P. E. (2014). Psychological strategies for winning a geopolitical forecasting tournament. *Psychological Science, 25*(5), 1106–1115. https://doi.org/10.1177/0956797614524255

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review, 115*(2), 502–517. https://doi.org/10.1037/0033-295X.115.2.502

Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing decisions: Improved decision making with a single training intervention. *Policy Insights from the Behavioral and Brain Sciences, 2*(1), 129–140. https://doi.org/10.1177/2372732215600886

Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin, 26*(9), 1142–1150. https://doi.org/10.1177/01461672002611010

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin, 28*(3), 369–381. https://doi.org/10.1177/0146167202286008

Sellier, A.-L., Scopelliti, I., & Morewedge, C. K. (2019). Debiasing training improves decision making in the field. *Psychological Science, 30*(9), 1371–1379. https://doi.org/10.1177/0956797619861429

Soll, J. B., Milkman, K. L., & Payne, J. W. (2015). A user's guide to debiasing. In G. Keren & G. Wu (Eds.), *The Wiley Blackwell handbook of judgment and decision making* (pp. 924–951). Wiley.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Whitesmith, M. (2020). *Cognitive bias in intelligence analysis: Testing the analysis of competing hypotheses method*. Edinburgh University Press.