Biases in decision making

ANTI-BIAS STRATEGIES By Tom Stafford

There are two kinds of bias typically studied by psychologists, both of which a judge will wish to avoid.

The first are the ‘social biases’, where we automatically form impressions of people, or leap to conclusions, based on the social group that they are a member of. Examples of social bias would be if we instantly warm to someone who speaks with the same accent as us, or if we assume someone from a different ethnic group is unlikely to be telling the truth.

The second kind of biases are ‘cognitive biases’, which are systematic tendencies in our thought processes that can lead us into error. The most famous is confirmation bias, whereby we seek information which can confirm our beliefs, inadequately testing beliefs by seeking out potentially contradictory information. Another cognitive bias is the ‘anchoring effect’, whereby, when making judgements about numerical quantities, we are overly influenced by the first number given to us (the anchor). So people who are first asked if Attila the Hun invaded Europe before or after 500 AD will give an earlier estimate when then asked for the exact year than people who are first asked if Attila the Hun invaded Europe before or after 1200 AD. The year in the initial before/after question anchors their subsequent estimate, distorting it in a similar way that shopkeepers hope that a price of £8.99 will make you think about an item as costing about eight pounds rather than the more accurate nine pounds.

Judicial bias?

Half a century of work by experimental psychologists has revealed two things about biases: We are all vulnerable to them, and they are difficult to account for, even if you know about them. Even for judges? Yes, even for judges.

One study found that 97% of US judges believe they are above the average in their ability to 'avoid racial prejudice in decision making', a statistically unlikely state of affairs which probably reflects judges overestimating their abilities with respect to those of their peers. Another study of white American judges found that they displayed the same automatic, 'implicit', negative associations with race as found in the general population.[1]

Finally, a study which asked judges to review genuine trial materials found that the sentences recommended were influenced by irrelevant suggestions – either when introduced by the suggestion of a prosecutor or by a probation officer. These irrelevant suggestions became 'anchors', demonstrating that judges – and judicial decisions – are prone to the anchoring effect, just like the rest of us.

Although there are few quick fixes for bias, there are diverse strategies which individuals and organisations can adopt...

Although there are few quick fixes for bias, there are diverse strategies which individuals and organisations can adopt which work against both cognitive and social biases. Many of these will already be recognised by working judges, or explicitly incorporated in legal procedures.

The requirement for written justification, reliance on objectively verifiable evidence, and even the adversarial system of prosecution and defence all play a role in preventing any individual from allowing their biases to run away with themselves.

Anti-bias strategies

Our work has focused on providing a framework to assist judges in thinking about their current anti-bias strategies, and about future bias strategies which they could adopt.

Our framework asks you to consider two dimensions on which any anti-bias strategies can be categorised. The first dimension is the locus of effect; we can divide anti-bias strategies by what their primary target is:

Personal strategies – which aim to change an individual's thoughts or behaviour
Interpersonal strategies – which target interactions between two or more people
Institutional strategies – which target the norms and regulations of the whole institution

We are often focused on the personal level – what can I do about bias, how can I avoid bias – but we should not forget that our work involves others, who will also have their own biases, and we can play an important role in addressing their biases, just as they can play a role in addressing ours (by holding us to account or asking for justification for potentially biased decisions).

Importantly, for anti-bias strategies to take hold they need to be moved beyond the level of individuals, so they are sustained by institutional support, not just individual effort (although of course the requirement for individual effort doesn't go away). Research shows that individuals often lack the perspective or resources to combat bias on their own, whilst successful and sustained change in outcomes requires institutional change.

The second dimension of our framework is the effect a strategy has on the bias.

Mitigation strategies – work against bias (but leave the bias intact)
Insulation strategies – remove the trigger for a bias, preventing it from occurring
Removal strategies – diminish the bias directly

Ridding ourselves of bias may be the best longer-term goal, but is likely to be slow and difficult. There is evidence that social biases born of ignorance, those which result in workplace discrimination on the grounds of sexuality, disability, or ethnicity, can be diminished in time by increasing workplace diversity.

Insulation strategies can be highly effective – for example university exam scripts are marked anonymously, so any prejudices towards students of different social groups are simply not triggered in those grading the scripts. Hiring panels are forbidden from asking certain questions (such as whether a job candidate hopes to get pregnant) and we can think of this as an insulation strategy. Insulation strategies have the drawback of not always being possible (for example you can't hide candidate gender during job interviews), and of leaving any potential bias unquestioned. They are important, but - like the other strategies – insufficient on their own.

1 Rachlinski, J. J., Johnson, S. L., Wistrich, A. J., & Guthrie, C. (2009). Does Unconscious Racial Bias Affect Trial Judges. *Notre Dame Law Review*, *84*, 1195–1246.

Mitigations strategies are perhaps the least likely to be effective, but easiest to immediately apply. This category includes everything from trying to avoid risk factors for bad decisions (like fatigue, hunger or being rushed) to systematic recording of decision outcomes so that any potential bias can be identified.

This gives us a 3 by 3 grid, which we can use to think about how we approach bias. What strategies do we already deploy, and where do they fit within the framework? Are there parts of the space which are under-populated, and could we think about adopting additional strategies there?

A 3x3 model

	Mitigate	Insulate	Remove
Personal	Avoid risk factors (hunger, fatigue), articulate reasoning, 'imagine the opposite'	Remove information that activates bias	Cognitive training (e.g. relearning associations)
Interpersonal	Identifying others' biases is easier; challenging conversations	Subdivide tasks to ensure independence of procedures; reveal identifying information last	Exposure to diversity ("Contact hypothesis")
Institutional	Tracking outcomes; predeclared criteria; recording process of decisions; norms of fairness	Procedures that remove bias activating information;	Avoiding biased outcomes (e.g. quotas / shortlisting requirements)

Shown here is the 3 by 3 grid populated with some examples of each type of strategy. It is neither incontestable nor exhaustive, but is intended to provide illustration of some of the strategies which are, or could, be adopted.

In conclusion, working against bias in our decisions is like healthy eating. You would not eat an apple and claim you had a healthy diet. Similarly, you cannot go on a bias awareness course and claim you now make unbiased decisions. Guarding against bias requires good habits, and good procedures. Effective anti-bias strategies need to be adopted for the long term by individuals, but also by the institutions within which we work.

Psychologists have been industrious in cataloguing the biases that can plague decision-making. The portrait of human rationality that has resulted is an overly pessimistic one. We can, and do, take effective action to reduce our biases.

For references to the studies mentioned in this piece and more information please see this website.

Tom Stafford is a Senior Lecturer in Psychology and Cognitive Science at the University of Sheffield

Back to contents