Debiasing the military appreciation process

Dr Jim Sheffield, Victoria University of Wellington
Colonel Roger Margetts, ONZM, New Zealand Army
Lieutenant Colonel Gordon Milward, ONZM, New Zealand Army

It ain’t what you don’t know that gets you into trouble. It’s what you know for sure that just ain’t so! (attributed, perhaps incorrectly, to Mark Twain)

Prior to the Iraq War, the Chairman of the US Joint Chiefs of Staff, Colin Powell, received advice about ‘weapons of mass destruction’ that turned out to be incorrect. He subsequently formulated three rules for his intelligence staff:

Tell me what you know. Tell me what you don’t know. And then, based on what you really know and what you really don’t know, tell me what you think is most likely to happen. (Powell, 2004)

Each of the rules encapsulates a distinct perspective or mode of thinking. In particular, the first two rules capture Mark Twain’s observation that what we know may not be so. Confidence in what we know comes fast and easy but must be balanced by the slow and difficult search for gaps in knowledge. The bias to confirm what we already know is strong. Even the hint of disconfirming evidence may cause negative emotion and ‘loss aversion’. Group norms may favour conformity over creative and critical thinking. A bias for information quantity may create complexity that obscures lack of information quality and the need for action.

In their article titled ‘Thinking more rationally: cognitive biases and the joint military appreciation process’, published in this Journal in 2015, Nigel Dobson-Keeffe and Warren Coaker asserted that ‘[t]here is little evidence that the Joint Military Appreciation Process (JMAP), as used by the ADF, considers the effects of these cognitive biases, despite their potential to affect operational decisions’. The authors then examined the JMAP to determine its vulnerability to bias, identifying what biases are important, and where they may have greatest impact.

However, we believe that complexity is a common reason why military decision-makers are unaware of the processes by which they make sense of the world, their own capabilities and
desires, and the interaction between them. We have focused, therefore, not on cognitive error as a deviation from a single, normatively correct response but on the confusions inherent in military decision-making. We collaborated with senior staff in the New Zealand Army to reflect critically on why the military appreciation process is important, and how it may be improved by a ‘de-biasing’ program.

**Literature review**

**System 1 and system 2 thinking**

The Military Appreciation Process (MAP) employed by the New Zealand Army is a structured decision-making process for operational planning primarily by individuals (where ‘groupthink’ bias may not be important). It is not dissimilar to the JMAP and other operational planning processes used by allied nations. Hoskin (2009) argues that the JMAP is generally very effective but could be improved by the better application of human factors, such as the personal role of the commander, intuition and creativity. In a similar vein, Smith (2011) argues that operational art and strategy are dynamic and contingent practices, and that expertise in the face of ambiguity requires an adaptive mental stance or cognitive dual-process capability.

Bloxam (2012) asserts that ‘in the last few decades there has been a gradual synthesis of the literature towards a dual-process model of cognition which is simple, but not simplistic, elegant and also very accessible’. While the terminology and application domains vary, there is consensus on the core concepts (see, for example, Evans, 2008; Kahneman, 2011). Evans & Stanovich (2013) state that:

> Our preferred theoretical approach is one in which rapid autonomous processes (Type 1) are assumed to yield default responses unless intervened on by distinctive higher order reasoning processes (Type 2). What defines the difference is that Type 2 processing supports hypothetical thinking and loads heavily on working memory.

These characteristics are illustrated at Table 1.

<table>
<thead>
<tr>
<th>Property</th>
<th>System 1 thinking</th>
<th>System 2 thinking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Action</td>
<td>Reflexive, skilled</td>
<td>Deliberate, rule-based</td>
</tr>
<tr>
<td>Automaticity</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Awareness</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Context important</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Cost</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Effort</td>
<td>Minimal</td>
<td>Considerable</td>
</tr>
<tr>
<td>Errors</td>
<td>Common</td>
<td>Few</td>
</tr>
<tr>
<td>Hard-wired</td>
<td>Maybe</td>
<td>No</td>
</tr>
<tr>
<td>Predictive power</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Reasoning style</td>
<td>Intuitive Heuristic Associative Concrete</td>
<td>Analytical</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Normative</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deductive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
</tr>
<tr>
<td>Reliability</td>
<td>Low, variable</td>
<td>High, consistent</td>
</tr>
<tr>
<td>Scientific rigour</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Speed</td>
<td>Fast</td>
<td>Slow</td>
</tr>
<tr>
<td>Vulnerability to bias</td>
<td>Yes</td>
<td>Less so</td>
</tr>
<tr>
<td>Working memory</td>
<td>Not involved</td>
<td>Involved</td>
</tr>
</tbody>
</table>

Table 1: Characteristics of system 1 and system 2 thinking (adapted from Bloxham, 2012)
Health

A key application domain of the cognitive dual-process model is health. Bloxham (2012) describes:

A common clinical sequence of events whereby the index case has an initial wrong diagnosis induced by its similarity to a more common or recently seen condition (availability heuristic), and which also appears to be a typical example of its class (representative heuristic). This initial diagnosis is then reinforced by several other well established heuristics (anchoring bias, confirmation bias and premature closure) which can all be contained under the scope of System 1. Treatment failure then forces a switch to System 2, which is more analytical in nature and may include the formal consideration of the combined probabilities of the various individual signs, symptoms and investigations pertaining to the case [as outlined at Figure 1].

Bloxham (2012) observes that both system 1 and system 2 thinking play an essential role in rational decision-making and that attention should be focused on the triggers or stimuli in the internal or external environment that promote or inhibit these two systems (Figure 2 and Table 2 refer).

Croskerry (2009) observes that the dual-process theory supports simplified models that can be readily taught to learners across a wide range of disciplines. In particular, he contends that

An understanding of the model allows for more focused metacognition, that is, the decision maker can identify which system they are currently using and determine the appropriateness and the relative benefits of remaining in that mode versus switching to the other.

Lambe, O’Reilly, Kelly & Curristan (2016) conducted a systematic survey of dual-process cognitive interventions to enhance diagnostic reasoning of medical trainees and doctors. While many of the studies found some effect of interventions, guided reflection interventions emerged as the most consistently successful. It was also concluded that cognitive forcing strategies improved accuracy and confidence judgments.
A second key application domain of the cognitive dual-process model is the military. All military decision-making articles addressed, implicitly or explicitly, fast (system 1) and slow (system 2) thinking, and the importance of the transition between them (see, for example, Ross, Evans, Dennis & Walsh, 2012). Aspects of the MAP aligned with system 1 are illustrated in Table 3.

1. Ill-defined goals and ill-structured tasks
2. Uncertainty, ambiguity, and missing data
3. Shifting and competing goals
4. Dynamic and continually changing conditions
5. Action-feedback loops (real-time reactions to changed conditions)
6. Time stress
7. High stakes
8. Multiple players
9. Organisational goals and norms
10. Experienced decision makers

Table 3: Aspects of the MAP aligned with system 1 thinking (source: Klein & Klinger, 1991)

Neuroscience is providing empirical evidence for the biology associated with the dual-process model, demonstrated in first-person accounts of combat experiences (Sheffield & Margetts, 2016). In debiasing, guided reflection about visceral experiences, and about emotions such as fear, would appear particularly useful.

In the Rand Corporation’s Project Airforce, Davis, Kulick & Egner (2005) recommended that ‘[d]ecision support should appeal to both the rational-analytic and the intuitive capabilities of the decisionmaker’. The authors contrasted the logic of storytelling (reduced conceptual uncertainty) and the logic of statistics (reduced empirical uncertainty), and stated that ‘the particular balance should depend on characteristics of the decision, the decision environment, and the decisionmaker’. The confirmation bias was a key barrier in moving from system 1 to system 2 thinking, with Cook & Smallman (2008) concluding that:

Domain experts exhibited confirmation bias in a realistic intelligence analysis task and apparently conflated evidence supportiveness with importance. Graphical evidence layout promoted more balanced and less biased evidence selection.

Wickens et al. (2010) found that when integrating intelligence information, anchoring on old (and less relevant) information sources (cues) was reduced by simple instructions to pay more attention to recent (more reliable) cues. Herdener
et al. (2016) found that when operators made trajectory predictions from graphical displays, they were able to update incorrect anchors (means) but were very poor at processing statistical information about uncertainty (variance).

In a report entitled ‘Training critical thinking for the battlefield’, Cohen et al. (2000) state that “[a] very short period of training has been consistently found to significantly affect both … variables related to critical thinking processes and … participants’ decisions in a military scenario’. Morewedge et al. (2015) report that a single training intervention with an instructional video or game produced large and persistent reductions in decision bias. Games which provided personalised feedback and practice produced larger effects than videos.

Models of the JMAP and MAP

The Dobson-Keeffe & Coaker (2015) representation of the inter-operable JMAP highlights that biases in collecting information (information, availability, pattern recognition) are important in three clusters of steps (labelled analysing data, deciding and acting). All four processes are marked by time pressures and conceptual uncertainty, so that cognitive processing is primarily experiential (Klein & Klinger, 1991; Slovic et al., 2004) or System 1 (Bloxham, 2012). The model makes predictions about which cluster (as opposed to which step) each bias will impact the most (Figure 3).

The current research builds on similarities between the battlespace intelligence operations of the ADF’s inter-operative JMAP and the New Zealand Army’s MAP. The Dobson-Keeffe & Coaker (2015) allocation of biases to the clusters of steps in the JMAP serves as a reference model, or initial best-practice theoretical model, in allocating biases to the clusters of steps in the MAP.

Our representation of the MAP highlights that biases in analysing information (confirmation, my side, illusory correlation, over-confidence) are important in three clusters of steps (labelled problem definition, design, and implementation). As the MAP is primarily viewed as an individual process, we have chosen to drop from our analysis the group-think bias. To sharpen the analysis, we have made predictions about which step (as opposed to which cluster of steps) each bias will impact the most. We envision a training program in which cognitive biases studied in the context of process flows serve as the occasion for possible transition between system 1 and system 2 (Figure 4).

Figure 3: JMAP reference model (adapted from Dobson-Keeffe & Coaker, 2015)
Figure 4: MAP theoretical model

Summary

In summary, the need for an appropriate balance between fast (system 1) and slow (system 2) thinking plays an important integrating role in the current behavioural decision-making literature. We reviewed applications and training programs in health and the military. Guided reflection, educational interventions, checklists, and cognitive forcing strategies were important tools. A single training intervention with an instructional game or video produced large and persistent reductions in decision bias.

A process model of the MAP was developed and predictions made about the biases important to each step. We argue that a training program focused on decision processes with embedded ‘metacognitive moments’ that serve as the occasion for possible transitions between system 1 and system 2 will enable more rational thinking in the New Zealand Army.

Methodology

The empirical research focuses on gaps between participant narratives and explores how a convergence of meaning can be fostered through a training program. Empirical data was gathered from an in-depth, qualitative case study of training for battlespace intelligence. Our informants (senior staff in the New Zealand Army) participated in a three-step research process (interviews, card sort and checklist).

Interviews

Semi-structured interviews guided the participants’ critical reflections on the purpose of the MAP, and its strengths and weaknesses, and how it could be improved (Margetts, 2016a).

Card sort

A card-sort exercise and a ‘think aloud’ protocol guided reflection on bias, and the likely impact on the eight steps in the MAP (Figure 5). Each bias in the Dobson-Keeffe & Coaker (2015) study was placed on an index card, with name and description on one side, and military application on the other. Observations were noted about the participants’ verbalisations, and their deliberations as they assigned biases to the steps. A bias could be allocated to more than one step, and not all biases had to be allocated. During this process, participants were encouraged to ‘think aloud’ about their experiences.

The allocation of biases to steps of the MAP was analysed against the criteria in three reference models:
1. The theoretical model created by transferring allocations to steps in the JMAP (Figure 3) to the steps in the MAP that arguably had a similar function (Figure 4).
2. An empirical model based on consensus on card-sort allocations.
3. A best-fit model based on an opportunistic choice of either the empirical or theoretical models.

For each model, patterns of participant bias assignments were analysed via 11 parameters of fit at four levels (including step and cluster) of abstraction.

Results

Interviews

Nine participants were involved in the interviews and card-sort activity, all with substantial military experience, to ensure they reflected the beliefs and opinions of individuals well qualified to understand and use the MAP. They also included a ‘key informant’, who had been central in the development of the MAP.

Why the MAP is important

The purpose of the MAP is to assist individuals in groups to plan and make decisions in a variety of demanding situations ranging from simple to complex, low stress to high stress and non-conflict right through to high-end conflict. It achieves this by aiding commanders and staff to think through what may happen and thereby be better prepared to effectively mitigate risks or exploit opportunities as they occur. The MAP is kind of doing two things; it’s assisting planning before execution and then assisting decision making during the execution of the plan. Whether the critical situation is time sensitive or not, whether it’s simple or complex, or whether you feel your life is threatened or not, it’s a tool with great utility. (key informant)

A strong theme was that the MAP ‘provides structure to the decision-making process’ and ‘breaks down complexity’. Positive comments such as ‘a method of interpreting complexity’ and ‘making sense of all the complexity’ recurred consistently. Negative comments included ‘leads people down rabbit holes’ and ‘not packaged user friendly’. Such respondents expressed reservations over the ‘time compression issues’ and preferences for ‘simpler tools’.

Overall, the usefulness of the MAP in managing complexity in critical situations was assessed at 4 on a 5-point scale (1=not at all useful; 5=extremely useful).

Strengths and weaknesses of the MAP

In the military context, we are often faced with a complex and dynamic situation that ideally requires slow thinking to solve, but in reality we have to react to using fast thinking. In these circumstances the MAP is very good at focusing a Commander and staff on determining when a decision has to be made by, and what type of planning approach is to be used. The MAP enables you to balance in the time available, the use of slow thinking analysis and fast thinking intuition. If you’ve got plenty of time you can run the full deliberate process and weight analysis over intuition—thereby avoiding the pitfall of using bias prone fast thinking to solve complex problems when you don’t have to. This exposes a potential weakness in military culture—because we are so used to using short term fast thinking in conflict situations we may tend to automatically apply this style of decision making to all situations. (key informant)

An emergent theme around weaknesses was seen in comments such as ‘at worst is inflexible and less responsive’ and ‘not supportive of quick, instinctive decision making’. It seems the deliberate nature of the process and the discipline it provides is also a drawback as it ‘limits divergent thinking’. Another theme centered on the frustration shared by students and instructors, such as it is ‘hard to teach—overwhelming for them and how to get into the detail they don’t understand’.

Common and recurring problems

It takes about six weeks for experienced military people to get up to speed on the process and then be able to intelligently focus on the situation. Learning or rather absorbing military culture, jargon and the lessons from operational experience is a socialisation process that you can’t speed up and takes between two to four years, in my opinion.
I think there’s a real danger of reinforcing only fast thinking as our default planning and decision making style. I think initially when you’re learning the process we need to take the time-sensitivity out so that people have the opportunity to practise slow thinking. There’s a real tension here, under pressure people just whip out the first answer that comes to mind which makes people even more prone to their bias. We could be inadvertently conditioning our officers to being decisive, unimaginative and bias prone decision-makers. (key informant)

For those new to the Army, the MAP is confounding as seen in the comment that ‘the new and inexperienced face a military culture … [and] a language they don’t understand’. Others capture a theme whereby the MAP ‘requires significant practice to master’, and again with the comment that ‘complexity and no prior experience [can] only be remedied by repetition of use’.

Goals for the training program

When using the MAP as an individual, challenging yourself to be a little bit creative in each of those steps that I talked about, and being aware of your own bias, is critical. Having some sort of mini process to be creative and avoiding being prone to bias, or whatever your frame of reference is, will be helpful.

At those points we should use some sort of tool or process that prompts us to think outside the box and avoid our instinctive and normal way of viewing things. That would help offset the danger of bias and lack of creativity. The key points of the MAP process that challenge your thinking become even more critical. So a smart checklist that mitigated bias at those critical points in the MAP that required assumptions and/or deductive reasoning would be very useful. (key informant)

A strong theme was that a checklist would promote a common awareness among instructors and junior officers of the impact of bias: ‘[a checklist would be] incredibly helpful [given we are] prone to bias, especially in initial use of the MAP’. Participants saw it as ‘definitely beneficial as an instructor’ as it gave ‘instructors [their] own understanding of bias’. A checklist was good ‘both ways—good for the student and instructors’. Overall, the usefulness of a debiasing checklist was assessed at 4.3 on a 5-point scale.

Interestingly, after experiencing the card-sort activity, some participants revised their rating upwards. No-one scored the checklist lower than 4. It appears that the card sort served to guide reflection and increase the flow of comments. A checklist would provide a focus when debriefing or correcting student performance. Also ‘[I’m a] big fan of that and good way to operate’ as it allows students to ‘look, think and act’. This appears to recognise that instructors and students would be able to share a common description and presentation of a bias, where it was likely to occur, and how to address it systematically.

Summary

Participants valued the MAP as a tool for managing complexity in critical situations. While it served many useful purposes, major issues remained, including irreducible complexity, insufficient training, the need for creativity, and a culture of time compression. There was ‘real danger of reinforcing only fast thinking as our default planning and decision making style’. Participants strongly believed that a debiasing checklist would address outstanding issues, reduce the impact of time compression, and promote balanced thinking. Participants strongly believed that a checklist and related debiasing procedures held positive benefit.

Card sort

Some indicative results of the card sort are presented in Figure 5. At the overall or MAP level, 182 assignments were made, with 81 assignments to the problem definition cluster, 59 to the design cluster, and 42 to the implementation cluster.

Some steps attracted more biases than others. For example, step 3 attracted 48 assignments, and the consensus was that biases 1, 2, 6, 7, and 10 (listed in Figure 5) were important. In contrast, step 1 attracted 7 assignments and there was a consensus that only bias 5 was important. The consensus allocation of biases differed in interesting ways from the reference model. For example, the consensus that biases 1, 2, and 10 have greatest impact at step 3 is at odds with the theoretical model, which requires that bias 1 and 2 be assigned to step 2, and bias 10 be assigned to step 8.
Debiasing the military appreciation process

Checklist

A debiasing checklist based on a cluster level analysis was developed as part of a 32-page curriculum pack. It was presented to the training staff and senior lecturers responsible for training the 2017 cohort of 30-40 junior officers. The training program, which included experiential exercises in metacognition and focused on guided reflection, was well received. The instructors also reported that teaching debiasing provided for their own reflection on the MAP, and the biases important to each step.

Summary

Empirical research employed a three-step methodology. Interpretive interviews with key decision-makers revealed that divergent beliefs existed and were associated with differing responses to time compression. Participation in a card-sort exercise induced MAP instructors and a key informant to believe that debiasing process flows would alleviate some of the frustrations occasioned by time compression.

Discussion

Two narratives were central to the interviews with participants, namely that ‘we could be inadvertently conditioning our officers to being decisive, unimaginative and bias prone decision-makers’; and that ‘a checklist would be good both for the students and instructors’. Participants strongly believed that a debiasing program would improve the MAP by mitigating the impact of time compression, and enabling a more appropriate balance of system 1 and system 2 thinking.

In the card sort, instructors’ cognitive explorations of the MAP were guided by multiple models, parameters, and levels of abstraction. The result was an empirical consensus allocation of biases that differed in interesting ways from the predictions of the theoretical model, indicating that further insights and accommodations are possible. Do the theoretical and empirical models identify real differences? Do our predictions at which step a bias is important make sense? What do those familiar with both the JMAP and the MAP think? Continuing discussions and exercises among ADF and New Zealand Army staff are expected to produce further insights and accommodation around interpretation.

Our analysis of decision processes in health and the MAP focused on individual decision-making. In health applications, the process model is somewhat linear, moving from diagnosis to treatment. Biases associated with system 1 tended to be located in the diagnosis cluster of steps, and biases associated with system 2

---

Figure 5: MAP empirical model (based on results of card sort)
in the treatment design and implementation clusters. The JMAP and MAP process models have steps and clusters that are richly interconnected, and highlight the continual need for information collecting and information analysis. This suggests that system 1 and system 2 thinking may be important in multiple steps and/or clusters. The checklist located the biases associated with information analysis (confirmation, my-side, illusory correlation, over-confidence) in all three clusters (problem definition, design, and implementation) of the MAP.

In summary, empirical research focused on participant narratives and fostered convergence of meaning through a training program that was seen as good both for students and instructors.

Conclusion

Our research with senior staff in the New Zealand Army identified that the root cause of different perceptions about the MAP was unresolved dilemmas about time compression. Convergence of sense and meaning was fostered by a training program that engaged MAP instructors, a key informant, and the current cohort of junior officers in guided reflection, a think-aloud protocol, card sorts, reference models, a checklist for debiasing the MAP, and metacognitive exercises.

The training program had a key takeaway message that was consistent with research within the sense-making paradigm and, we believe, is applicable to many project-based organisations:

Don’t beat yourself up (about bias) as we need to make assumptions to survive everyday life. As you eat you assume that the food is not poisoned, although you have no evidence. As you drive through a green light, you assume the cross traffic will remain stopped, although you have no proof. The key is to have an awareness of our assumptions and realise that sometimes our intuition can hurt us and lead to a bad assumption gap. Whether consciously or subconsciously, once you make an assumption, the gap between what you think and the reality can get further and further apart. (Margetts, 2016b, p. 6)

This message was perhaps one reason why the checklist was perceived as good ‘both ways’. Acknowledgement of bias had become an occasion for creativity and an accepted way of enabling more rational thinking in the New Zealand Army. Previous research indicates that similar one-shot, relatively brief training programs were also received with some enthusiasm and had lasting effect. The research therefore illustrates a practical and cost-effective approach to developing debiasing training programs.

A focus on cognitive dual-process capability (system 1 and system 2) and the transitions between them has enabled the design of a practical debiasing program. This program integrated many biases under a single theoretical perspective and empirical process model. Future research will explore related applications in the context of the New Zealand Army, and allied forces including the ADF.

Dr Jim Sheffield is Director of Case Competitions at Victoria University of Wellington, and Senior Lecturer, School of Management, where he teaches problem solving and decision making to MBA students. He first researched heuristics and biases at the University of Arizona in the 1980s. His doctoral work in group support systems led to the creation of a Decision Support Centre at the University of Auckland. Major action research initiatives include the development of economic strategy, science policy and comprehensive regional planning. He has over 100 publications.

Colonel Roger Margetts is presently serving in Iraq as Director of Training for the Iraqi Security Forces. He has served in the New Zealand Defence Force since 1989, including command roles in the 1st Infantry Battalion, the 1st NZSAS Regiment and the Land Operations Training Centre. Roger has served extensively in the Middle East through multiple tours in Kuwait, Afghanistan and Iraq. Key staff leader roles include Head of Capability Management Teams and Instructor Counter Insurgency at the Australian Command and Staff College. Roger has a MBA from Victoria University of Wellington, and a Bachelor of Arts degree in Psychology from Massey University.

Lieutenant Colonel Gordon Milward is the Program Principal Lead Capability at the New Zealand Defence Force’s (NZDF) Institute for Leader Development. He has commanded the 2/1st Infantry Battalion, the 3rd Land Force Group, the NZ Army Training Group and the NZDF Land Component. Other roles have been Chief of Staff Operations in East Timor, Chief of the Observer Group Golan in the Middle East, and the NZ Defence Advisor to PNG, Solomon Islands and Vanuatu. In 1996-97, while on the Directing Staff of the Australian Army Command and Staff College, he developed and trialled the planning and decision-making
Debiasing the military appreciation process

Debiasing the military appreciation process, the genesis of the current Joint Military Appreciation Process.

Acknowledgements: The authors wish to thank Lieutenant Colonel Andrew Fox and staff at the New Zealand Army Command School for their generous support for this research project.

References

Davis, P.K., J. Kulick and M. Egner (2005), Implications of modern decision science for military decision-support systems, RAND Corporation: Santa Monica.
Hoskin, R. (2009), The ghost in the machine: better application of human factors to enhance the military appreciation process, Land Warfare Studies Centre: Canberra.