Bilalić, M., McLeod, P., & Gobet, F. (in press). The mechanism of the Einstellung (set) effect: A pervasive source of cognitive bias. *Current Directions of Psychological Science*.

COVER PAGE

Correspondence address:

Merim Bilalić

Tübingen University

Experimental MR

Department of Neuroradiology

Hoppe-Seyler-Str. 3

72076 Tübingen, Germany

Phone: +49 7071 2987 384

Fax: +49 7071 294 371

Email: merim.bilalic@med.uni-tuebingen.de

Word count: 2,344 (2,538 with the Box Future Directions)

The Mechanism of the Einstellung (Set) Effect:

A Pervasive Source of Cognitive Bias

Merim Bilalić[1]

Tübingen University

Peter McLeod

Oxford University

and

Fernand Gobet

Brunel University

Abstract

The eye movements of expert players trying to solve a chess problem show that the first idea that comes to mind directs attention towards sources of information consistent with itself and away from inconsistent information. This bias continues unconsciously even when the player believes he is looking for alternatives. The result is that alternatives to the first idea are ignored. This mechanism for biasing attention ensures a speedy response in familiar situations but it can lead to errors when the first thought that comes to mind is not appropriate. We propose that this mechanism is the source of many cognitive biases from phenomena in problem solving and reasoning, to perceptual errors and failures in memory.

Key words – cognitive bias; Einstellung (set) effect; problem solving; chess; expertise

The Mechanism of the Einstellung (Set) Effect: A Pervasive Source of Cognitive Bias

Four centuries ago, Francis Bacon pointed out the unfortunate human tendency to ignore new evidence which could undermine a firmly held opinion (Bacon, 1620/1939; p. 36.) Little has changed. Tetlock (2005) found that expert political scientists do not change their theories when events prove their predictions wrong; they keep the theories and discount the evidence. Similarly Gould (2006) showed that scientists can be so strongly influenced by the theory they already hold that they do not interpret new data objectively. The experts' theories were originally based on an accumulation of evidence, so it is not that they cannot absorb new information. The question is: Why, once a point of view has been formed, do people find it difficult to assimilate new information if it is not consistent with the view already held?

The answer is suggested by recent studies of the Einstellung (mental set) effect – the fixation of thought produced by prior experience - which demonstrate its power and reveal its mechanism (Bilalić, McLeod, & Gobet, 2008a; 2008b). Many cognitive biases that make it difficult for people to assimilate new evidence may have their origin in a similar mechanism to that which produces the Einstellung effect. The difficulties people have in assimilating new evidence may have the origin in a similar mechanism to that which produces the Einstellung effect. The effect is particularly dangerous because people are unaware that they are affected by it. As Stephen J. Gould noted: "In most cases … biases … were unknowingly influential and ... scientists believed they were pursuing unsullied truth" (p. 59).

THE CLASSICAL EINSTELLUNG (SET) EFFECT

The Einstellung effect occurs when an idea that comes immediately to mind in a familiar context prevents alternatives being considered. It was first experimentally demonstrated by Luchins (1942) with the water-jug problems. He gave people a series of

problems that could be solved by a fixed method which they quickly learnt. Then he gave them a problem that could be solved using the usual method but also with a different, quicker one (called the 2-solution problem because of two possible solutions). Most of the participants continued to use the old method, not spotting the quicker alternative. This is perhaps unsurprising given that the old method solved the problem. However, when participants were presented with a final problem (called the 1-solution problem because there was only one solution possible), apparently similar to the previous ones but where the familiar method now did not work, many of the participants said that this problem was insoluble. In fact, it could be solved using the shorter method from the previous problem. The fixation of thought displayed by these people was demonstrated by a control group who were given only the 1-solution problem. They solved it quickly, showing that the problem was not intrinsically difficult. The experimental group failed to find the solution because the similarity of the final problem to the previous ones brought the usual (but now inappropriate) method to mind, blinding them to alternatives.

THE EINSTELLUNG EFFECT IN (CHESS) EXPERTS

We recently showed that the Einstellung phenomenon can not only be demonstrated with laypeople and simple logical problems but also with experts and complex problems (Bilalić, McLeod, & Gobet, 2008b). We showed the position in Figure 1A to expert chess players and asked them to find the shortest solution. This is a 2-solution problem just as in the Luchins study. One solution is a five move sequence leading to a smothered mate. The smothered mate sequence is well known to all good chess players and the possibility that it could be used in this position was quickly noticed by all the players. The second solution is less familiar but shorter, leading to mate in three moves. Players spoke aloud as they tried to solve the problem. All players found the familiar solution quickly and then said they were looking for

a shorter one. Those who failed to find it (and said that the smothered mate was the shortest solution) were then shown the 1-solution problem (Figure 1B). This problem is the same as the 2-solution problem except that the familiar smothered mate solution has been disabled by moving one piece, leaving only the shorter solution from the 2-solution problem. Hence, the 1-solution problem was similar to the last problem in the Luchins design where the participants' familiar solution would not solve the problem but an alternative method would. All players found the shorter solution in the 1-solution problem, showing that they were capable of finding it when not distracted by the familiar one[2].

QUANTIFYING THE EINSTELLUNG EFFECT

Previously the Einstellung effect has only been reported as a qualitative effect. With chess players it can be quantified as the relative strength of different players is known precisely from their performance other players of known strength. The scale has a mean of 1500 and a standard deviation of 200. Table 1 gives the names of different groups and their strength relative to average players in terms of the number of standard deviations they are above the average). The Einstellung effect was quantified by seeing how much weaker a player had to be, when only the shorter solution was present (the 1-solution problem), to show comparable performance to that of a better player with the 2-solution problem when the distracting effect of the familiar solution was present. Across a range of skill levels, the presence of a familiar solution that first came to mind reduced the problem solving performance of the experts to that of players about three standard deviations lower in skill (see Table 1). The Einstellung effect is indeed a very powerful effect – the chance of a player being beaten by a player three standard deviations lower in skill is close to zero. Yet that is the level to which the first idea that came to mind on seeing the position – the presence of a smothered mate – reduced the ability of the players to find another solution.

Table 1. 2-solution problem (see Figure 1A for the problem). The percentage of stronger players who found the less familiar, shorter solution in the presence of the more familiar but longer solution. Player strength is given in SDs above the average for all chess players. *1-solution problem* (see Figure 1B for the problem). The percentage of both stronger and weaker players who found the less familiar, shorter solution when the more familiar one had been removed. Weaker players were presented with the 1-solution problem only. The difference between the 2-solution and 1-solution problem is the presence of the familiar solution which produces the Einstellung effect. We quantified the effect by comparing the performance of stronger players on the 2-solution problem with the performance of weaker players on the 1-solution problem. The performance of International Masters (5 SDs above average) on the 2-solution problem was comparable to that of Class A players (2 SDs above average) on the 1-solution problem, the performance of Masters (4 SDs above average) was comparable to that of the Class B players (1 SD above average) and the performance of Candidate Masters (3 SDs above average) was the same as Class C players (average). Experiments with different problems and the more natural instruction to find the best solution rather than the shortest, yield similar quantitative differences between the performance of stronger players on the 2-solution version of the problems and the performance of the weaker players on the 1-solution version of the problems (see Bilalić, McLeod, & Gobet, 2008b).

	Problem	
Skill Level	*2-solution*	*1-solution*
Stronger players		
International Master (+5 SD)	50%	100%
Master (+4 SD)	18%	100%

Candidate Master (+3 SD)	0%	100%
Weaker players		
Class A (+2 SD)		63%
Class B (+1 SD)		13%
Class C (average)		0%

THE MECHANISM OF THE EINSTELLUNG (SET) EFFECT

Why did the players experiencing the Einstellung effect fail to find the less familiar solution? We measured the eye movements of two new groups of chess experts who were given either the 2-solution problem or the 1-solution problem (Bilalić, McLeod, & Gobet, 2008a). This allowed us to see which squares the players were looking at and how long they spent looking at them as they tried to solve the problem. We used Candidate Masters (3 SD above average) and as in the previous experiment none of the group shown the 2-solution problem (Figure 1A) found the shorter solution while all the players shown the 1-solution problem (Figure 1B) found it. Although all the players trying to solve the 2-solution problem said that they looked for a shorter solution after spotting the familiar smothered mate solution, the eye movements, shown in Figure 1C, told a different story. The players' eyes continued to dwell on squares and pieces involved in the familiar smothered mate solution (as shown by the green circles) throughout the time they believed they were looking for alternatives. They spent little time on the squares required to find the shorter solution (as shown by the red triangles). The group shown the 1-solution problem found the shorter solution without much difficulty. Initially, their attention, as measured by their eye movements shown in Figure 1D, was directed equally at squares and pieces involved in the solution and at those that were not. But shortly

before announcing that they had discovered the solution they started to focus on the key squares.

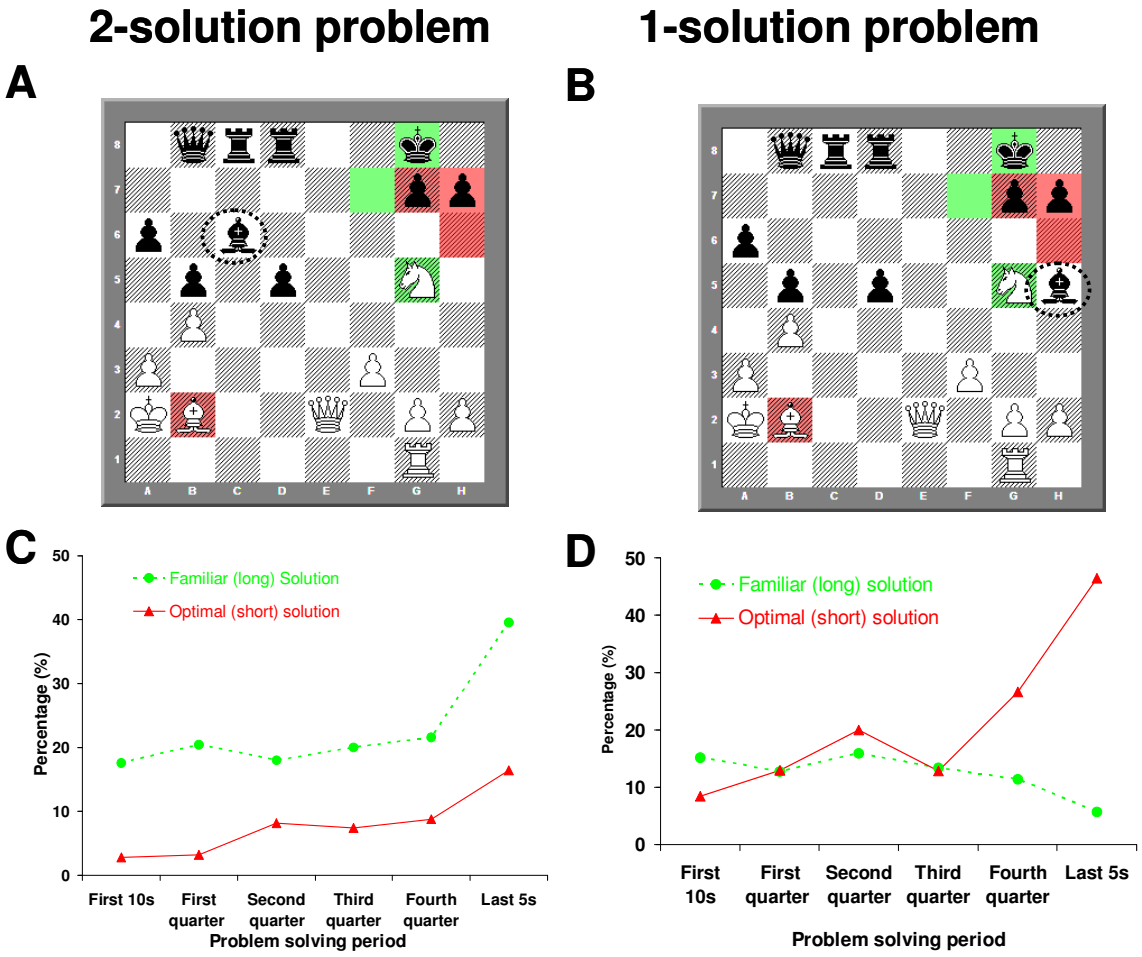

Figure 1.

The eye movement data demonstrate how a pattern of thought, once activated, can prevent other patterns of thought becoming active. As soon as a problem or situation is recognised as familiar, the knowledge (schema) for dealing with it is activated. The schema directs attention towards those aspects of the situation that are relevant to itself and away from those that are not. Thus a self-fulfilling circle begins with information consistent with the already activated schema being more likely to be picked up and inconsistent information ignored. Consequently the belief that the schema is the right one to deal with the situation is

confirmed and alternatives are less likely to be considered. At a conscious level the individuals think they are considering the evidence in an open-minded way, unaware that their attention is being directed selectively to certain aspects of the task. Those things that they notice are consistent with the active schema and they are confirmed in the view that the way they are dealing with the situation is the correct one. Things that do not fit in are either not noticed, or if they are, discarded because they do not fit the activated schema.

A PERVASIVE BIASSING INFLUENCE ON COGNITION

Problem solving failures caused by the Einstellung effect are the downside of a normally efficient cognitive mechanism. We rely on the knowledge acquired through by past experience to deal quickly with the familiar. It seems inefficient to spend time looking for an alternative solution if we already have an adequate one. Indeed, in complex real world situations people usually prefer to look for solutions that are good enough rather than trying for an elusive best that may be out of reach (Simon, 1990). Good solutions come from previous experience. But sometimes, as the Einstellung phenomena shows, this may be disadvantageous.

We believe that a similar mechanism may lie behind a range of biases that have been shown in everyday thought. Once someone has a firmly entrenched idea about politics, the character of a colleague or the best way to perform a task, it can be difficult to persuade them to think differently (Gardner, 2004; Rokeach, 1960). People will accept a lower standard of evidence when accepting evidence that supports their view than they will do it for evidence that goes against it (Lord, Ross, & Lepper, 1979). People look for evidence that will confirm a currently held view rather than evidence that might disconfirm it (Wason, 1960; for a review, see Nickerson, 1998). The notoriously difficult insight problems studied by (Duncker, 1945) present an extreme case of the inability to overcome previously activated schemas. People repeatedly tried to solve the problem with the same method although it had repeatedly proved

unsuccessful. Even constant failure to find a solution is not enough to prevent people thinking of the same method when they face the same problem again (Knoblich, Ohlsson, & Raney, 2001).

Expertise does not prevent this bias. Doctors form opinions quickly based on previous experience, often missing important aspects which are inconsistent with their initial opinion (Groopman, 2007). Political experts and scientists are so heavily influenced by their favourite theories that they ignore valid negative evidence (Gould, 1996; Tetlock, 2005). Experts do not realize that their favoured view seems so good because their attention has been directed to information that supports it and away from information that does not.

A similar mechanism may lie behind biases in other areas of cognition. The part-set cuing phenomenon in memory demonstrates the distracting effect of already activated knowledge. For example, people who are asked to recall the names of American states and given a number of state names as examples, recall fewer names than people who are not given the examples (Brown, 1968; see also Anderson, Bjork, & Bjork, 1994). The memories which have been activated by the experimenter impede access to other, un-activated areas of memory. The biasing of perceptual interpretation by prior knowledge is demonstrated when a well-learned schema overwrites perceptual input. For example, correct description of a playing card requires a much longer exposure if the colour is reversed (a black three of hearts) than if it is normal. The effect can be surprisingly powerful. Even with exposures of a second, many cards are reported in their conventional rather than real colouring (a black three of hearts reported as a red three of hearts; Bruner & Postman, 1949).

In each case described above, already activated knowledge biased the way people subsequently perceived and interacted with the world. This bias from schemas developed from previous experience is a blessing - without it we would have to deal with every situation as if we were encountering it for the first time. The mechanism which produces the Einstellung effect and many other biases in cognition shows that it can also be a curse. As John Maynard

Keynes once said, “The difficulty lies, not in the new ideas, but in escaping from the old ones, which ramify ... into every corner of our minds” (1936/1973; p. xxiii).

BOX: FUTURE DIRECTIONS

The Einstellung effect resembles many biasing phenomena both in the laboratory and in everyday life. We hope that future work will establish if the mechanism we have shown to be behind one version of the Einstellung effect – the first activated thought biasing the subsequent allocation of attention and perceptual input – is also responsible for the other related phenomena. This ambitious goal may be approached with a mixture of behavioral and neuroimaging techniques. For example, the recordings of eye movements together with think aloud protocols during the classical Luchins water jug problem might show why people are unable to solve the 1-solution problem. Similarly, the mixture of eye tracking and think aloud protocols with the paradigms involving the perceptual judgements, confirmation bias and reasoning may provide evidence on the mechanism behind these phenomena. Neuroimaging techniques have been applied to understanding the brain mechanisms behind phenomena of selective attention which show some similarities to the biases in thought and memory discussed here. Future neuroimaging studies of people experiencing Einstellung may show that the same control circuits are involved in the selective biasing of thought and memory retrieval as are involved in selective pick-up of sensory information.

Notes

[1] Address correspondence to Merim Bilalić, Tübingen University, Experimental MR, Department of Neuroradiology, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany; e-mail: merim.bilalic@med.uni-tuebingen.de

[2] Unlike Luchins' participants, all expert players found the solution once the familiar solution had been removed in the 1-solution problem. The effect had a cost on the problem solving process as shown by another group of experts who solved the 1-solution problem immediately, without being exposed to the 2-soution problem. This group of experts, which was comparable in skill with the one which solved both problems, found the optimal solution in half the time (37 seconds) it took the group who had previously experienced the Einstellung effect (78 seconds). This result is at first sight paradoxical. The players who were exposed to the 2-solution problem before were more familiar with the problem but nevertheless were slower to find the solution. The result is explained by a constant influence of the Einstellung effect which continues to distract experts even when it is removed (see Bilalić, McLeod, & Gobet, 2008a).

References

Anderson, C. M., Bjork, A. R., & Bjork, L. E. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1063-1087.

Bacon, F. (1939). Novum organum. In Burtt, E. A. (Ed.), *The English philosophers from Bacon to Mill* (pp. 24-123). New York: Random House. (Original work published in 1620)

Bilalić, M., McLeod, P., & Gobet, F. (2008a). Why good thoughts block better ones: The mechanism of the pernicious Einstellung effect. *Cognition*, *108*, 652-661.

Bilalić, M., McLeod, P., & Gobet, F. (2008b). Inflexibility of experts – Reality or myth? Quantifying the Einstellung effect in chess masters. *Cognitive Psychology, 56,* 73–102.

Brown, J. (1968). Reciprocal facilitation and impairment in free recall. *Psychonomic Science*, *10*, 41–42.

Bruner, J. S., & Postman, L. (1949). On the perception of incongruity: A paradigm. *Journal of Personality*, *18*, 206-223.

Duncker, K. (1945). On problem-solving. *Psychological Monographs, 58, ix,* 113.

Gardner, H. (2004). *Changing minds*. Cambridge, MA: Harvard Business School Press.

Gould, S. J. (1996). *The mismeasure of man*. (2nd ed). New York: W. W. Norton & Company.

Groopman, J. (2007). *How doctors think.* New York: Houghton Mifflin.

Keynes, J. M. (1973). *The general theory of employment, interest and money*. London: MacMillan, London. (Original work published 1936)

Knoblich, G., Ohlsson, S., & Raney, E. G. (2001). An eye movement study of insight problem solving. *Memory & Cognition, 29*, 1000-1009.

Lord, G. C., Ross, L., & Lepper, R. M. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*, 2098-2109.

Luchins, A. S. (1942). Mechanization in problem solving - The effect of Einstellung. *Psychological Monographs, 54*, No. 6, 95.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2,* 175-220.

Rokeach, M. (1960). *The open and closed mind.* Oxford, England: Basic Books.

Saariluoma, P. (1990). Apperception and restructuring in chess players' problem solving. In K. J. Gilhooly, M. T. G. Keane, R. H. Logie, & G. Erdos (Eds.), *Lines of thought: Reflections on the psychology of thinking* (pp. 41-57). London: Wiley.

Simon, H. A. (1990). Alternative visions of rationality. In P. K. Moser (Ed.), *Rationality in action: Contemporary approaches* (pp. 189-204). New York: Cambridge University Press.

Tetlock, P. (2005). *Expert political judgement: How good is it? How can we know?* Princeton, NJ: Princeton University Press.

Wason, C. P. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12*, 129-140.

Figure captions

Figure 1. **A)** The 2-solution problem. Players were instructed to find the shortest way for White to win the game. There are two solutions. The longer smothered mate solution is: *1*. Qe6+ Kh8 *2*. Nf7+ Kg8 *3*. Nh6++ Kh8 *4*. Qg8+ Rxg8 *5*. Nf7 mate. A less familiar shorter solution is: *1*. Qe6+ Kh8 *2*. Qh6! Rd7 *3*. Qxh7 mate, or *2*.… Kg8 *3*. Qxg7 mate. **B)** The 1-solution problem. This is the same position as the 2-solution problem except that the Bishop (encircled by the black dots) is on h5 rather than c6. Smothered mate is now no longer possible because Black's bishop covers f7. The shorter solution is still possible. *1*. Qe6+ Kh8 (If *1*..... Kf8 *2* Nxh7 mate) *2*. Qh6! Rd7 *3*. Qxh7 mate, or *2*.… Kg8 *3*. Qxg7 mate, or *2* … Bg6 *3*. Qxg7 mate. The squares crucial for the longer solution but not for the shorter one are highlighted in green (f7, g8 & g5) and those crucial for the shorter solution but not the longer one in red (b2, h6, h7 & g7). The problems are based on an idea by Saariluoma (1990). **C)** The percentage of time in different phases of problem solving that the experts (Candidate Masters) solving the 2-solution problem spent looking at squares crucial to the familiar long solution (green color) and shorter solution (red color). For each player the first 10 s and the last 5 s before announcing the solution were analysed separately. The remaining period, of whatever length, was divided into four equal periods. **D)** The same graph but this time showing the percentage of time the experts looking at the 1-solution problem spent looking at the crucial squares for the familiar longer and shorter solution. The familiar solution was disabled here.

Recommended Readings

[1] Luchins, A. S. (1942). (See References). A historical classic; introduces the Einstellung (mental set) phenomenon in the psychological literature. [Can be obtained at the web collection of classical contributions in psychology *http://psychclassics.yorku.ca/*]

[2] Bilalić, M., McLeod, P., & Gobet, F. (2008b). (See References). Shows that the Einstellung effect can be demonstrated with experts, quantifies its strength, and offers theoretical explanations of the phenomenon.

[3] Bilalić, M., McLeod, P., & Gobet, F. (2008a). (See References). Demonstrates the mechanism behind the Einstellung effect.

[4] Samuel S, Kundel H. L, Nodine C. F., & Toto L. C. (1995). Mechanism of satisfaction of search: eye position recordings in the reading of chest radiographs. *Radiology, 194,* 895-902. The paper explains why expert radiographers tend to miss other important abnormalities once they have already found an abnormality.

[5] Takahashi K., & Watanabe K. (2008). Persisting effect of prior experience of change blindness. *Perception, 37*, 324-7. A witty study showing that even cognitive scientists are not spared the negative influence of prior knowledge in a change detection task.