Cognition 108 (2008) 652–661

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect

Merim Bilalić [a,*], Peter McLeod [a], Fernand Gobet [b]

[a] Oxford University, Department of Experimental Psychology, South Parks Road, OX1 3UD, Oxford, UK
[b] Brunel University, School of Social Sciences, Uxbridge, UB8 3PH, UK

ARTICLE INFO

Article history:
Received 5 February 2007
Revised 6 May 2008
Acepted 8 May 2008

Keywords:
Einstellung (set) effect
Fixation
Problem solving
Expertise
Eye movements
Chess
Confirmation bias
Thinking

ABSTRACT

The Einstellung (set) effect occurs when the first idea that comes to mind, triggered by familiar features of a problem, prevents a better solution being found. It has been shown to affect both people facing novel problems and experts within their field of expertise. We show that it works by influencing mechanisms that determine what information is attended to. Having found one solution, expert chess players reported that they were looking for a better one. But their eye movements showed that they continued to look at features of the problem related to the solution they had already thought of. The mechanism which allows the first schema activated by familiar aspects of a problem to control the subsequent direction of attention may contribute to a wide range of biases both in everyday and expert thought – from confirmation bias in hypothesis testing to the tendency of scientists to ignore results that do not fit their favoured theories.

1. Introduction

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

The counter-intuitive possibility that prior knowledge can have a negative effect on future performance is a theme in a range of areas of psychology that at first sight might seem unrelated. For example, in negative transfer paradigms previous experience makes it more difficult to adapt to a new setting than it would be without such experience (Besnard & Cacitti, 2005; McCloy, Beaman, Morgan, & Speed, 2007; Schwartz, Bransford, & Sears, 2005; Singley & Anderson, 1989; Thorndike & Woodworth, 1901). Similarly, in the part-set cueing phenomenon people who are given a portion of previously studied/memorized material were in fact hampered by this 'help' when they tried to recall the remaining material compared to people who were just asked to recall the material (Basden & Basden, 1995; Roediger, 1973. Another example is offered by insight problems, first investigated by Köhler (1917/1925), Maier (1930), Maier (1931), and Duncker (1945). Although there are different theories of why these problems are so hard to solve, most agree that the difficulty lies in the initial mental representation of the problem, determined by prior knowledge, from which people are unable to escape even though it does not lead to a successful solution (Kaplan & Simon, 1990; Kershaw & Ohlsson, 2004; Knoblich, Ohlsson, Haider, & Rhenius, 1999; MacGregor, Ormerod, & Chronicle, 2001; Ohlsson, 1992; Smith, 1995). A similar effect can be found in the motor system. Common actions become automated through frequent use and are triggered by familiar situations without conscious thought. This usually leads to efficient responses to the demands of every-

* Corresponding author. Address: Department of Neuroradiology, Experimental MRI, Tübingen University, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany, Tel.: +49 7071 2987 384; fax: +49 7071 294 371.
E-mail addresses: merim.bilalic@med.uni-tuebingen.de, merim.bilalic@googlemail.com (M. Bilalić).

0010-0277/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2008.05.005

day life. But when a specific context is strongly associated with a well-learnt, but now inappropriate, pattern, the familiar pattern can emerge with unintended results (Norman, 1981). These may be amusing (or embarrassing) if one dials a familiar telephone number at the wrong time (e.g., Woltz, Gardner, & Bell, 2000). But the result can be catastrophic if someone controlling a nuclear reactor executes a common but now inappropriate sequence (Reason, 1990). Finally, the way people seek information is biased by their prior knowledge. When people test a theory, they look for evidence that is consistent with what they already believe rather than objectively assessing any evidence even if it might disconfirm their previously held belief (Dawson, Gilovich, & Regan, 2002; McKenzie, 2006; Nickerson, 1998; Wason, 1960; Westen, Kilts, Blagov, Harenski, & Hamann, 2006).

An instance of the negative impact of previous knowledge is known as the Einstellung (set) effect. This occurs when the first idea that comes to mind, triggered by previous experience with similar situations, prevents alternatives being considered. If this initial idea is not the best way to solve the problem, the optimum solution may be missed. The effect was first demonstrated experimentally by Luchins (1942), who 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 appeared similar to the previous ones but which could not be solved by the same method (the 'extinction' problem). Many said it was insoluble. The fixation of thought displayed by these people was demonstrated by a control group who were given only the extinction 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 (and now inappropriate) method to mind, preventing them from considering alternatives.

The Einstellung effect has been repeatedly demonstrated in the laboratory in a variety of forms using problems that do not require prior knowledge of the problem domain (e.g., Atwood and Polson, 1976; Chen and Mo, 2004; Delaney, Ericsson, and Knowles, 2004; Lippman, 1996; Lovett and Anderson, 1996). These studies show that mental set can be induced by a small number of similar problems in people who have never experienced the task before. It can also be used to demonstrate the mechanisms behind the mistakes made by experts. Experts rarely make mistakes but when they do, it is usually because they think the situation is a familiar one and apply their usual, but now inappropriate, methods to find a solution (Reason, 1990; Singley & Anderson, 1989). For example, most errors that doctors make are not connected to their inadequate medical knowledge but rather to the tendency to form opinions quickly based on previous experience. Once the initial diagnosis is formed, it guides doctors in the search of supporting evidence which in turn brings dangers of missing important aspects unrelated to the initial diagnosis (Croskerry, 2003; Graber, Franklin, & Gordon, 2005; Groopman, 2007; Kassirer & Kopelman, 1989). Such errors by experts with real life problems (where the errors are induced by knowledge of a well-learned procedure in long-term memory) resemble errors induced by Einstellung problems in the laboratory (where the mind set is induced by short-term memory of a newly learned procedure).

Despite the potential importance of the Einstellung effect for understanding why experts sometimes fail to find the optimum solution to a problem within their area of expertise, there are few studies investigating the effect with experts (for similar research on how experts can fail *outside* their domain, see and Wiley, 1998; Frensch & Sternberg, 1989; Hecht & Proffitt, 1995). One set of studies that has looked at this is Saariluoma's (1990), Saariluoma's (1992) demonstration that skilful chess players can fall victim to the Einstellung effect. The players tried to solve a number of chess problems where a familiar motif led to the solution. They were then presented with a problem that had two solutions. One was the familiar solution they had used to solve the previous problem, but was sub-optimal; the other was less familiar but optimal. Most players failed to find the optimal solution in the presence of the familiar solution. Saariluoma concluded that chess players can suffer from the Einstellung effect.

Bilalić, McLeod, and Gobet (2008) extended Saariluoma's observations, confirming that expert chess players can experience the Einstellung effect, and showing that it can be quantified. They showed expert players (Candidate Masters, Masters, and International Masters[1]) a number of problems such as the one shown in Fig. 1a, and asked them to find the shortest way for White to achieve checkmate. There were two solutions, one a well-known solution (smothered mate) taking five moves and the other (the optimal solution) a less familiar one taking three moves. The players who found the familiar solution but failed to find the optimal one were then shown a similar problem where the familiar solution had been disabled, leaving only the optimal one (shown in Fig. 1b). All the experts found the shorter solution in the 1-solution problem, showing that they were capable of finding it when not distracted by the familiar one. To quantify the impact which the familiar solution had on the performance of expert players in the 2-solution problem, Bilalić et al. (2008) exploited one of the advantages of chess as a domain for studying problem solving – the presence of an interval scale for measurement of skill (see Footnote 1).

Similarly to the studies on transfer, Bilalić et al. (2008) used a control group of weaker players (Class A, Class B, and Class C) and showed them the 1-solution problem only. The Einstellung effect was quantified by seeing how much weaker a player had to be, when only the optimal solution was present, to show comparable performance to that of a better player when the distracting effect of the familiar solution was present (for similar measures of transfer see the first chapter of Singley & Anderson, 1989). The performance of the International Masters (5 SDs above average) on the 2-solution problem was

[1] Chess skill is measured with the Elo rating, an interval scale with a theoretical mean of 1500 and a theoretical standard deviation of 200 (Elo, 1978). Average players with the rating between 1400 and 1600 are called Class C players, 1600–1800 Class B (+1SD above the mean), 1800–2000 Class A (+2SD), 2000–2200 are called Candidate Masters (+3SD), 2200–2400 Masters (+4SD), 2400–2500 International Masters (around 5SD above the mean), and finally, players above 2500 are called Grand Masters.