A nice bit of scandal:
About a disconfirmation bias in the Wason's 2-4-6 problem

Jean-Paul Caverni & Sandrine Rossi

Centre National de la Recherche Scientifique
et Université de Provence à Aix-en-Provence, France

Abstract

A well-known finding in human reasoning is the so-called "confirmation bias": When testing hypotheses, human beings try to confirm rather than disconfirm them. The aim of this paper is to reveal a possible artifact in the way data have been usually coded in the Wason's 2-4-6 problem. The existence of such an artefact leads to the unexpected and quite provocative conclusion that people do not exhibit a confirmation bias, as commonly believed, but exhibit instead a disconfirmation bias.

Rather than comparing the current triple with the present hypothesis, as is usually done, we compared it with the previous hypothesis proposed by participants. Every time a triple is a negative example of the previous hypothesis it can be assumed that participants used a disconfirmatory strategy. This phenomenon is shown to occur in up to 70% of the cas

Correspondence should be sent to Jean-Paul Caverni, Centre de Recherche en Psychologie Cognitive, CREPCO, CNRS & Université de Provence (UMR 6561), 29 Avenue Robert Schuman, F-13621 Aix-en-Provence Cedex 1, France (E-mail: caverni@newsup.univ-mrs.fr).

Vittorio Girotto and Patrick Lemaire made many helpful comments on a earlier version of this article, for which we are grateful.

A well-known finding in human reasoning is the so-called "confirmation bias" (see Evans, 1989): When testing hypotheses, people try to confirm instead of disconfirm them. This phenomenon has been shown in several studies (for a comprehensive review see Evans, 1989; Garnham & Oakhill, 1994; Gorman, 1995). One of the most well-known tasks in which it has been widely documented is the so-called "2-4-6 problem" (Wason, 1960). The aim of this short paper is to reveal a possible artifact in the way the data have been usually coded. Because of this artifact one could conclude that people exhibit not a confirmation bias but a disconfirmation bias.

Let us first recall the task. Participants are told that the experimenter has a rule in mind that classifies number triples and they have to discover this rule. They are given an example of a triple satisfying the to-be discovered rule, namely 2-4-6. In order to find the rule, participants have to generate triples of their own. For each triple they generate, they are told whether it does or does not satisfy the to-be discovered rule. Participants keep a written record of the triples they proposed, their current hypothesis for each of these, as well as the experimenter's feedback. They are told not to state the rule until they are sure that it is the correct one. If the stated rule is not the to be-discovered rule, participants continue to test additional triples.

In general, participants propose triples that are positive examples of the current hypothesis they claim they are testing (e.g. the "ascending with equal interval" hypothesis is tested with triples such as "10-12-14", "10-20-30", etc.). The prevailing interpretation (e.g. Wason, 1960; Tweney et al., 1980) has been that participants were prone to confirmation bias: They would rather verify (confirm) their hypotheses (giving positive examples) than falsify (disconfirm) them (giving negative or counter examples).

Because the to be-discovered rule is "any ascending sequence", when giving triples as "10-12-14" or "10-20-30" participants are told that these triples satisfy the rule and receive a non-conclusive verification of their (wrong) hypothesis (such as "numbers increasing in intervals of two", "increasing even numbers", "increasing multiples of the first number", "arithmetic progression", etc.). A good triple to receive a conclusive refutation would be a negative example of the tested hypothesis, as for instance"10-20-50" for the "ascending with equal interval" hypothesis.

We want to examine the following question: Is a negative example of the current hypothesis the one and only way to reject this hypothesis? Let us consider the following protocol (Table 1) in which the participant first tests a "even numbers" hypothesis by a "4-6-8" triple and then a "odd numbers" hypothesis by a "3-5-7" triple. As in both cases the participant tests her/his hypotheses by proposing examples rather than counter-examples, it can be argued that he/she is exhibiting a confirmatory strategy. However, such an interpretation can be criticised. In particular there are no theoretical arguments to assume that people test their hypotheses (1) by considering only the current triple they give, and (2) by failing to take into consideration both the previous and present hypotheses.

<Insert Table 1 about here>

It can be argued that in the second step of the indicated sample the participant's goal is to disconfirm the previous "even number" hypothesis. To this aim, he/she uses a positive example of the "odd number" alternative, which is a negative example (a counter-example) of the "even number" hypothesis. Given that he/she will be informed that both triples satisfy the rule, he/she will be able to reject both properties "even" and "odd" as belonging to the to-be-discovered rule. Thus, one might conclude that the participant exhibited not a confirming but a disconfirming strategy.

Tracing such a strategy involves comparing the current triple not with the current hypothesis (as has been the usual procedure), but with the previous hypothesis that the participant has proposed (i.e. the hypothesis he/she proposed on the preceding trial n-1). Every time the current triple is a negative example of the previous hypothesis, it can be assumed that participants use a disconfirmatory strategy.

Is such a strategy used in solving the 2-4-6 problem? In order to answer this question we ran an experiment in which participants ($\underline{n}$= 67) were presented with the classic instructions of 2-4-6 problem (cf. Wason, 1960). Participants were graduate and undergraduate students (in various disciplines, excluding psychology) at the Faculty of Literature and Humanities at the University of Provence in Aix-en-Provence, France.

For each participant, we computed the proportion of confirmations of the total number of triples, by comparing each triple both with its current hypothesis (classical coding) and the hypothesis held for the triple considered immediately prior (alternative coding). Table 2 gives a sample of protocol.

<Insert Table 2 about here>

Figure 1 gives the mean percentages for these two comparisons, considering first the whole sample of participants (89.46 vs 48.76; $F(1,66)=138.40$, $\underline{p}<.0001$), and then separately those ($\underline{n}$=56) who found the to-be-discovered rule (88.52 vs 45.30; $F(1,55)=154.13$, $\underline{p}<.0001$). Among this group 18 participants found the rule at their first announcement (89.17 vs 33.78; $F(1,17)=87.83$, $\underline{p}<.0001$). In all cases the mean for the classical coding using the current hypothesis was significantly higher than the mean for the alternative coding using the previous hypothesis. For the successful participants who found the rule at the first announcement this mean was 33.78%: in other words 66.22% of their hypothesis tests were disconfirmatory! The same pattern of results was for the 11 remaining unsuccessful participants (94.27 vs 66.36; $F(1,10)=6.36$, $p<.03$). In addition, the mean for unsuccessful participants was significantly greater than the mean for the successful participants at their first attempt (66.36 vs 33.78; $F(1,27) = 9.33$, $p<.005$) and the overall mean for successful participants (66.36 vs 45.30; $F(1,65) = 5.73$, $p<.02$).

<Insert Figure 1 about here>

We ran a final analysis, comparing among the successful participants those who found the rule at their first attempt and those who did not, again using both coding methods. Although the means were not significantly different for the classical coding (89.17 vs 88.21; F<1), there was a significant difference when using the alternative coding method considering previous hypothesis in favour of successful participants who were correct on their first attempt (33.78 vs 50.76; $F(1,54)=6.51$, $p<.01$).

Discussion

In the first two decades of life of 2-4-6 problem, the prevaling interpretation of the results was that participants are prone to a confirmation bias. According to Evans (1983) however, the phenomenon arises not from a deliberate motivation to avoid falsification but from a cognitive limitation in processing negative information: "It is not that subjects do not wish to falsify, it is simply that they cannot think the way to do it " (p.143). For Evans people’s difficulty can be better described as resulting from a positivity bias than from a confirmation bias. Moreover, Klayman and Ha (1987) claimed that testing a hypothesis with a positive example should not be considered as a confirmation strategy. Indeed, if the rule to be discovered is more specific than the hypothesis tested, the latter can only be disconfirmed by positive examples which are compatible with the hypotheses but not with the rule to be discovered. For example, if the rule to be discovered would be "three consecutive even numbers" and the participant test the hypothesis "three increasing even numbers" using the triple 4-8-10, then the experimenter’s answer would be "no" and would induce the participant to disconfirm her/his hypothesis. According to Klayman and Ha, it is because of the particular relationship between the subject’s hypothesis (specific) and the rule (general) that this positive strategy turns out to be ineffective in the usual 2-4-6 problem, whereas it is effective in most real life problems.

Our analysis showed that this mechanism, the positive test strategy, is also effective in the 2-4-6 problem, as in such a problem it also led to the rejection of hypotheses. It seems likely that in our experiment, participants were able to test a hypothesis and its alternative. However they did so with two successive trials, whereas it would have been usual to expect them to do so with only a single trial. According to Evans (1983), this former positive strategy is probably easier than the latter. In addition, according to Klayman and Ha (1987), it is as efficient as the latter. To test both a hypothesis and its alternative within the same trial, participants should be told that only a few number of trials are allowed. Alternatively they could receive feedback at the same time both on their current hypothesis and the alternative, as in the Tweney's et al. (1980) DAX-MED problem. In this experiment, participants were told some triples were MEDs and others DAXs. Their performance improved dramatically: 60% solved the DAX rule on the first announcement. For Tweney and his colleagues, when participants were told that a triple was a MED one, they would produce a negative version of their MED hypothesis as a DAX hypothesis and test both.

Whatever the case, our protocol analysis showed that most participants used a confirmatory as well as a disconfirmatory strategy, and that the successful participants preferred to use a disconfirmatory strategy. In most cases, they did so by giving a positive example of their current hypothesis and receiving positive feedback. We propose that this way of testing hypotheses reflects not a confirmatory strategy, as has been assumed to date, but instead reflects a disconfirmatory strategy.

Such a provocative conclusion inevitably raises some questions. (1) Why do participants change their hypothesis when they have received positive feedback to its regard? As Klayman and Ha (op.cit.) pointed out, positive feedback does not allow consideration of whetehr an hypothesis is true or not. The way to determine a positively answered hypothesis is true is to test its alternative. (2) Does the change of hypothesis express a participant's aim to disconfirm or is it merely an exploration of her/his space problem? As most disconfirmatory participants were also the most successul, we can infer that by and large the hypothesis change expresses a deliberate strategy.

References

Evans, J. St B. T. (1983). Selective processes in reasoning. In J. St B. T. Evans (Ed.), *Thinking and reasoning: Psychological approaches.* London: Routledge and Kegan Paul.

Evans, J. St B. T. (1989). *Bias in human reasoning.* Hove and London (UK): Lawrence Erlbaum Associates.

Garnham, A., & Oakhill, J. (1994). *Thinking and reasoning.* Oxford (UK) & Cambridge (USA): Blackwell.

Gorman, M.E. (1995). Hypothesis testing. In S.E. Newstead & J. St B. T. Evans (Eds), *Perspectives on Thinking and Reasoning.* Hove (UK) & Hillsdale (USA): Lawrence Erlbaum.

Klayman, J., & Ha, Y-W. (1987). Confirmation, disconfirmation and information in hypothesis testing. *Psychological Review, 94,* 211-228.

Tweney, R.D., Doherty, M.E., Warner, W.J., & Pliske, D.B. (1980). Strategies of rule discovery in an inference task. *Quaterly Journal of Experimenal Psychology, 32,* 109-124.

Wason, P.C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12,* 129-137.

Table 1. Theoretic sample protocol

triple	tested hypotheses
4-6-8	even numbers
3-5-7	odd numbers

Table 2. Sample protocol from our experiment

Subject n°20

triples proposed	tested hypotheses	feedback	classical coding (1)	alternative coding (2)
4-6-8	Even numbers increasing by two	YES	C	/
6-8-10	Even numbers increasing by two	YES	C	C
10-14-18	Increasing even numbers	YES	C	D
14-18-10	Even numbers independant of order and interval	NO	C	D
120-122-128	Increasing even numbers	YES	C	D
rule proposed: Even increasing numbers independant of interval				
18-14-10	Even numbers decreasing by four	NO	C	D
18-10-06	Decreasing even numbers	NO	C	D
18-16-14	Even numbers decreasing by two	NO	C	D
14-16-18	Even numbers increasing by two	YES	C	C
14-18-22	Even numbers increasing by four	YES	C	D
01-02-03	Increasing even and odd numbers	YES	C	D
rule proposed: Increasing numbers				

(1) Classical coding takes into account the triple and the hypothesis at the same step (e.g. step n°2: 6-8-10; even numbers increasing by two: confirmation).

(2) The proposed alternative coding takes into account the triple and the previous hypothesis only at previous step (e.g. step n°3: 10-14-18; even numbers increasing by two: disconfirmation).

C = Confirmation
D = Disconfirmation

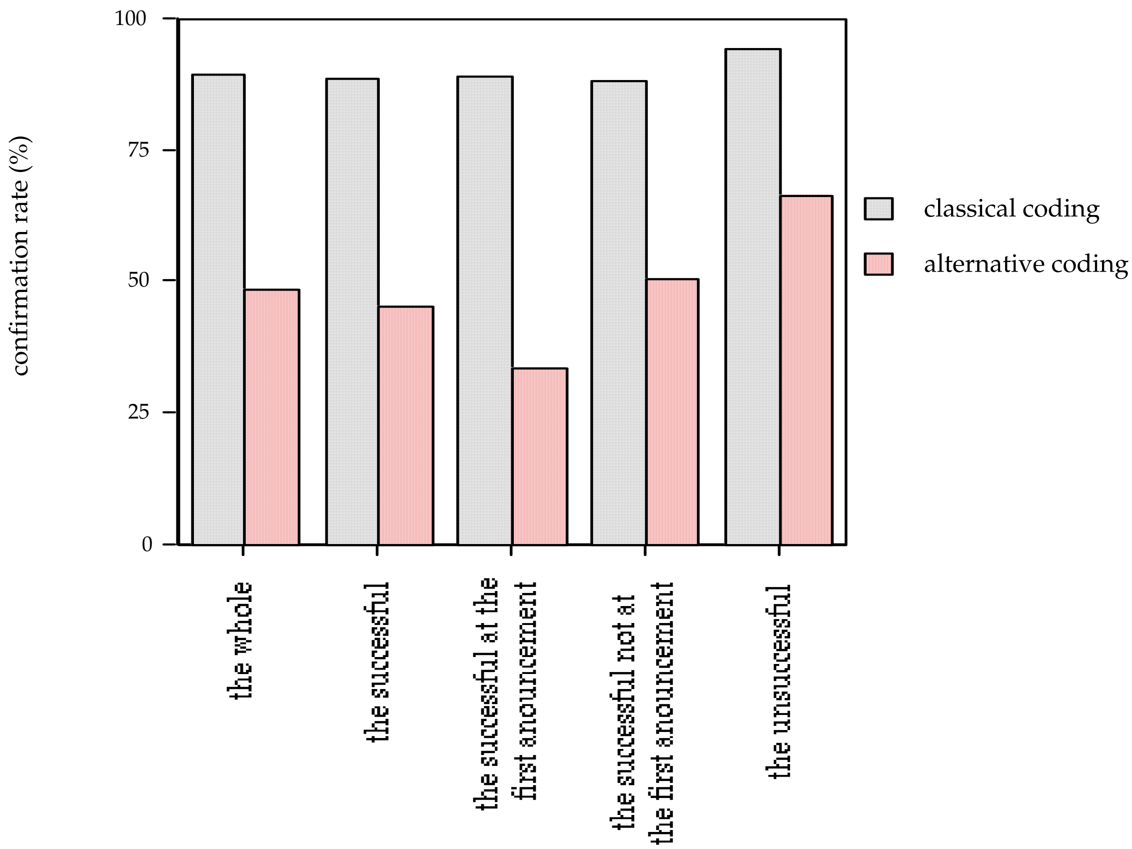

Figure 1. Confirmation rates observed for the both types of coding, indicated for the all types of subjects (the whole, the successful, etc.)