On a generality of confirmation bias: Individual differences perspective

Predrag Teovanović (teovanovic@fasper.bg.ac.rs)
Faculty for Special Education and Rehabilitation, University of Belgrade, Serbia
Laboratory for Research of Individual Differences, University of Belgrade, Serbia

Vincent Berthet (vincent.berthet@univ-lorraine.fr)
Psychology and Neuroscience Lab, Université de Lorraine, France
Centre d'Économie de la Sorbonne, France

Abstract

Although confirmation bias is being described as the most prominent cognitive bias, it is still unclear if it should be conceptualized as a unitary construct, or rather as an umbrella term for the group of related but independent phenomena. In the present study, we collected nine measures of individual differences in susceptibility to confirmation bias in different aspects of cognitive processing by employing and modifying several experimental paradigms. Results showed an overall preference for confirming over disconfirming information, but indicated that some cognitive tasks need further improvements in order to obtain more reliable measures. Nevertheless, the appliance of confirmatory factor analytic framework for testing hierarchically structured constructs showed that confirmation biases in information search, the weighing of evidence and memory recall are relatively independent constructs. Correlations among them were fair but insufficient to indicate the presence of a reliable second-order confirmation bias factor. In sum, the observed pattern of correlations suggests that it is more plausible to conceptualize confirmation bias as a multifaceted construct.

Keywords: *confirmation bias; individual differences; selection task; 2-4-6 task; interviewee's personality task*

Introduction

Confirmation bias refers to situations in which "information is searched, interpreted, and remembered in such a way that it systematically impedes the possibility that the hypothesis will be rejected" (Oswald & Grosjean, 2004, p. 79). It has been described as the best-known cognitive bias that came out of the literature on human reasoning (Evans, 1989).

Confirmation bias is observed in a wide range of real-life contexts and experimental paradigms, as well as in distinct aspects of cognitive processing, which led some authors to argue that there exists a variety of discrete confirmation biases rather than one unified phenomenon (see, e.g., Klayman, 1995; Nickerson, 1998). Motivated by this notion, we have employed the methodological approach of differential psychology to examine if various measures of individual differences in confirmation bias indicate that it is better to understand it as a unitary construct, or as a set of relatively related but still independent phenomena. For this purpose, we have included three experimental paradigms in our study, and adopted them for measuring individual differences in susceptibility to confirmation bias in three distinct cognitive processes. Such an approach is not usual in the field, considering that most of the confirmation bias studies are focused on one aspect of information processing and one paradigm (for a partial exception, see Vedejová & Čavojová, 2022).

Paradigms that are regularly used to demonstrate confirmation bias include the four-card selection task (Wason, 1960), the 2-4-6 task (Wason, 1966) and the interviewee's personality task (Snyder & Swann, 1978). What is common to these tasks is an instruction given to participants to test a hypothesis by searching for specific information. For example, participants might be asked to indicate which cards among E, C, 4 and 9 they would select to examine the rule "If a card has a vowel on one side, then it has an even number on the other side", or to indicate which questions they would ask to check if someone is emotionally stable. A typical finding is that people far more often look for evidence that might confirm/support a given hypothesis (e.g., selecting card "5", or asking the question "What do you do to keep calm in pressure situations?") than for evidence that might falsify/disapprove a given hypothesis (e.g., selecting card "7", or asking the question "How do you show your nervousness?").

Confirmation bias has also been detected in other cognitive processes such as the weighing of evidence. For example, previous research demonstrated that participants evaluate incongruent information much more critically than congruent information (Lord et al., 1979), and opt to retain the initial hypothesis when faced with either ambiguous (Hoch & Ha, 1986) or falsifying evidence (Beattie & Baron, 1988). In our study, we asked participants to indicate for each piece of evidence, either confirming (e.g., "on the back of card 2 is an A") or disconfirming (e.g., "on the back of card 3 is a U"), whether it was informative in deciding if the given rule is true or false.

Finally, some studies demonstrate confirmation bias in memory recall by showing that confirming information is more often perceived as previously encountered than disconfirming information (see, e.g., Eagly et al., 1999; Stangor & Mcmillan, 1992). In the current study, we presented several outcomes to participants, some of which had already been presented in previous tasks, while the others were completely new, and asked them to indicate which outcomes were presented before.

In sum, we employed three experimental paradigms (selection card task, 2-4-6 task and the interviewee's personality task), and modified them to measure confirmation bias in three cognitive processes - information search (IS), the weighing of evidence (WE) and memory recall (MR). We used a total of nine cognitive tasks to measure individual differences in confirmation bias to compare the strength of its different manifestations, and to explore if the pattern of their correlations suggests that it is more plausible to conceptualize confirmation bias as a singular or multifaceted phenomenon.

Method

Participants

Tasks were administered to the community sample (*N* = 200; 41.5% women) who were recruited by the Fédération S2CH in France to complete an online session, for which they received compensation of €10. The mean age of the participants was 27.08 (*SD* = 4.73). The research was reviewed and approved by the Institutional Review Board–Paris School of Economics (approval number: 2020-022).

Measures

As previously indicated, we have employed a total of nine cognitive tasks in order to measure individual differences in confirmation bias. All tasks included confirming and disconfirming information, and some of them also included neutral information. Participants were asked to respond by indicating which information they would choose in order to examine the given hypothesis (IS tasks), which evidence they find informative (WE tasks), or which evidence had already been presented (MR tasks). On each task, the confirmation bias score was calculated as a difference between responses to confirming and disconfirming information. A more detailed description of the procedures as well as all materials used and the data collected in the present study, are available at https://osf.io/saj4c/.

Procedure

After providing consent, participants completed the tasks in the following order: information search, the weighing of evidence and memory recall. Within each block, participants were first presented with the selection card task, then the 2-4-6 task, and finally the interviewee's personality task.

Analytical strategy

After examining the experimental reliability of confirmation bias across nine cognitive tasks, correlational data were analyzed. To extract more meaningful information from the correlation matrix and to directly test the proposed research question, we have applied a confirmatory factor analytic framework for testing hierarchically structured constructs (Brunner et al., 2012). We examined five competing measurement models. The first assumes that only one general factor underlies performance on all nine tasks (one-factor model). The second model presumes that the performance on nine tasks is organized under three more specific paradigm factors (selection task, 2-4-6 and the interviewee's personality), while the third model presumes that organizing specific factors are three distinct cognitive processes (information search, the weighing of evidence and memory recall). Both second and third measurement models had two versions in which either (a) correlated but relatively independent specific factors were postulated (first-order factor models without general higher-order factor), or (b) correlations between specific first-order factors were restricted to zero, and uncorrelated general factor that directly influences all nine measures was added to the model (bifactor/nested-factor model that assumes hierarchical structure).

Results

On the whole, participants responded by choosing a confirming option in 69.1% of cases (SD = 13.6), and a disconfirming option in 48.1% of cases (SD = 15.0). This difference was highly significant ($t(199) = 12.61$, $p < .001$), showing a strong overall confirmation bias ($d = 0.89$, 95% CI [0.73, 1.05]).

As indicated by the results presented in Table 1, confirmation bias was detected on seven out of nine tasks ($ps < .001$), on which the mean effect of the confirmative information was between 0.27 and 1.04 standard deviations. However, this tendency was not significant on two out of nine tasks in our study, both within the interviewee's personality paradigm, one concerning the weighing of evidence ($p = .22$) and the other concerning memory recall ($p = .70$).

Table 1. Differences between confirming and disconfirming choices on nine tasks

Task	Confirming M (SD)	Disconfirming M (SD)	$t(199)$	p	Cohen's d
Selection Card IS	64.1 (38.3)	21.8 (31.5)	10.90	< .001	0.77
Selection Card WE	80.0 (29.4)	62.9 (38.5)	4.66	< .001	0.33
Selection Card MR	60.0 (24.0)	43.2 (25.1)	5.88	< .001	0.42
2-4-6 IS	87.3 (27.3)	29.8 (36.7)	14.65	< .001	1.04
2-4-6 WE	77.7 (32.8)	64.8 (32.1)	3.78	< .001	0.27
2-4-6 MR	65.3 (24.1)	42.2 (20.1)	10.38	< .001	0.72
Interviewee IS	53.8 (17.1)	36.8 (18.3)	7.26	< .001	0.51
Interviewee WE	84.5 (22.5)	82.2 (23.4)	1.23	.22	0.09
Interviewee MR	49.4 (17.7)	49.9 (16.8)	-0.40	.70	-0.02

Note. IS – Information Search, WE – Weighing of Evidence, MR – Memory Recall

Table 2. Reliabilities (on diagonal) and correlations among nine confirmation bias scores

Measure (number of items)	1	2	3	4	5	6	7	8	9
1. Selection Card IS (4)	.84								
2. Selection Card WE (4)	.23*	.93							
3. Selection Card MR (4)	.12	.48**	.86						
4. 2-4-6 IS (3)	.16	.42**	.28**	.75					
5. 2-4-6 WE (3)	.06	.66**	.25**	.38**	.68				
6. 2-4-6 MR (3)	-.10	.27**	.34**	.20*	.38**	.63			
7. Interviewee IS (4)	.13	.26**	.13	.30**	.28**	.08	.64		
8. Interviewee WE (4)	.18*	.43**	.35**	.20*	.29**	.15	.17	.69	
9. Interviewee MR (4)	-.05	.10	.28**	.05	.04	.13	.05	.33**	.48

Note. IS – Information Search, WE – Weighing of Evidence, MR – Memory Recall; *$p < .01$, **$p < .001$

Table 3. Fit indices for five competing models

Model	χ^2	*df*	*p*	CFI	RMSEA	AIC
1. One-factor model	65.11	26	< .001	.89	.09 [.06 - .11]	3918
2a. First-order paradigm factors	61.48	23	< .001	.89	.09 [.06 - .12]	3920
2b. Nested-factor paradigm factor	59.99	17	< .001	.88	.11 [.08 - .14]	3930
3a. First-order process factors	39.57	23	.02	.95	.06 [.03 - .09]	3898
3b. Nested-factor process factor	32.53	17	.01	.96	.07 [.03 - .10]	3901

The confirmation effect observed on the information search task ($t(199) = 16.74$, $p < .001$, $d = 1.18$, 95% CI [1.00, 1.36]) was significantly larger than the corresponding effect on the memory recall task ($t(199) = 8.34$, $p < .001$, $d = 0.59$, 95% CI [0.44, 0.74]) and the weighing of evidence task ($t(199) = 4.36$, $p < .001$, $d = 0.31$, 95% CI [0.17, 0.45]).

On the other hand, the effect observed within the interviewee's personality paradigm ($t(199) = 5.16$, $p < .001$, $d = 0.36$, 95% CI [0.22, 0.51]) was significantly smaller than the confirmation effect within the selection card paradigm ($t(199) = 10.19$, $p < .001$, $d = 0.72$, 95% CI [0.56, 0.88]), the 2-4-6 task paradigm ($t(199) = 12.82$, $p < .001$, $d = 0.91$, 95% CI [0.74, 1.07]).

However, even when mean differences were not observed, there was considerable variability in the responses of participants, as indicated by the standard deviations of bias scores.

An acceptable level of internal consistency of individual differences across items was observed consistently only on selection card tasks, which is in line with some of the previous research (see, e.g., Stanovich & West, 1998), while the noisiest measures of individual differences were detected on the interviewee's personality tasks (see Table 2).

Correlations between nine confirmation bias scores were modest on average, but they ranged from zero to as much as .66. The mean correlation was .22, while 24 out of 36 correlations were statistically significant (*ps* < .05).

In the next step, we tested five competing CFA measurement models. Results presented in Table 3 indicate that the model with three correlated but independent first-order cognitive processes factors fitted the data ($\chi^2(23) = 39.57$, $p = .02$; $CFI = .95$, $RMSEA = .06$, 90% CI [.03 - .09], $AIC = 3898$; see Figure 1) significantly better than the one-factor model ($\Delta\chi^2(3) = 25.54$, $p < .001$). Contrary to that, the inclusion of three experimental paradigms factors did not improve the fit of the model significantly in comparison to the one-factor model ($\Delta\chi^2(3) = 3.62$, $p = .31$). On the other side, attempts to further improve the fit of the model by assuming hierarchical structure, that is, by including three specific as well as one general factor, were not successful, either in the case of paradigm ($\Delta\chi^2(6) = 1.49$, $p = .96$) or in process-centered measurement models ($\Delta\chi^2(6) = 7.04$, $p = .31$).

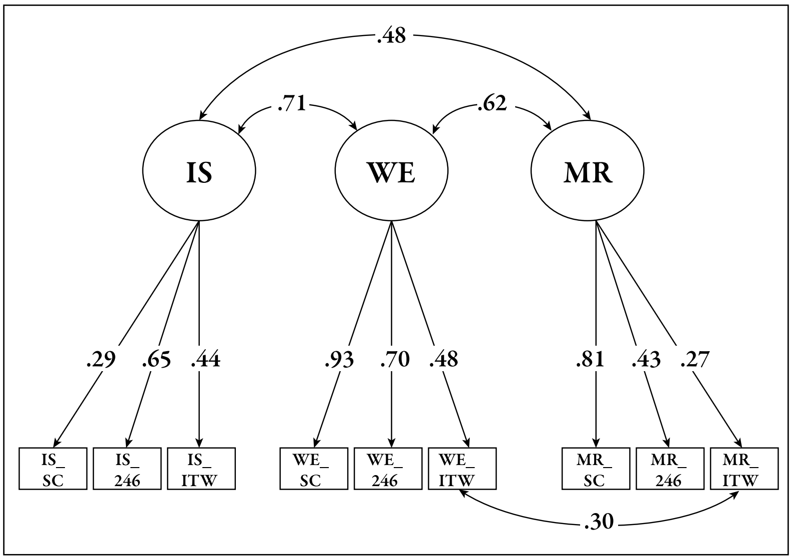

Figure 1. First-Order Factor Model with Three Related Process Factors[1]

In sum, the most parsimonious way to organize various measures of individual differences in susceptibility to confirmation bias was to assume three first-order cognitive processes factors that correlate fairly but still insufficiently to indicate the presence of a reliable second-order general confirmation bias factor. Although related, confirmation biases in IS, WE and MR should be seen as relatively independent constructs.

Discussion

Confirmation bias in the present study was observed on a wide range of tasks, and it was relatively robust, which confirms previous findings within the field (Evans, 1989). Following the conceptual distinction introduced by Oswald and Grosjean (2004), it has been shown that some confirmation bias guises are more deceptive than others, and that distortions are especially pronounced in the case of information search. However, reasons for this might be at least partially methodological in the sense that employed tasks were originally designed to measure biased IS (Snyder & Swann, 1978; Wason, 1960, 1966), while tasks intended to measure the biased WE and MR were subsequently derived from them.

The observed reliability of individual differences was the highest in the case of all three selection card tasks, which is in line with some of the previous findings (Stanovich & West, 1998). Also, results concerning both experimental and differential reliability indices show that the interviewee's personality paradigm alternations of WE and MR tasks were the least successful. This could be due to the specific semantic content of items that were highly loaded with meaning. Furthermore, it has been suggested that differences in the extent to which proposed questions on this task are actually diagnostic with respect to testing given hypotheses might also explain participants' choices (Trope & Bassok, 1983). In that respect, future studies that seek to employ the interviewee's personality paradigm for the purpose of measuring proneness to confirmation bias might benefit from using more abstract or at least content-controlled versions of the task.

The strategy of using multiple measures of confirmation bias within a single study is relatively rare in the field. Vedejová and Čavojová (2022) have recently examined confirmation strategies of hypothesis testing in three aspects of cognitive processing (IS, WE and MR) by using a unified paradigm. Their study showed that bias was most prevalent in IS, followed by WE, while they fail to detect a biased MR.

[1] It should be noted that the relation between the measures of confirmation bias on two interviewee's personality tasks, namely the weighing of evidence and memory recall, was introduced since it was indicated by the high values of modification indices in all tested models (>10). Nevertheless, this did not influence the main conclusions of the study.

However, we opted to include a wide range of cognitive tasks not only to compare observed effect sizes of confirming information across different tasks, but also to collect a variety of measures of individual differences in susceptibility to confirmation bias and study the pattern of correlations between them. This question brings theoretical importance, considering the stance that confirmation bias is not a single unitary construct, but rather a set of relatively related but still independent phenomena (Klayman, 1995; Nickerson, 1998). For this purpose, we examined five competing measurement models by using a confirmatory factor analytic framework for testing hierarchically structured constructs (Brunner et al., 2012). Results indicated that observed heterogeneity in correlations among different manifestations of confirmation bias stems from the differences between distinct aspects of cognitive processing (IS, WE and MR), rather than from the differences between different experimental paradigms (selection card, 2-4-6 and the interviewee's personality tasks). Furthermore, measurement models which included a general confirmation bias factor (assuming either its sole existence or its overarching role above first-order specific factors) were significantly inferior in comparison to the measurement model, which included only three related but independent first-order factors of confirmation bias in different aspects of cognitive processing. Put differently, the observed pattern of correlations suggests that it is more plausible to conceptualize confirmation bias as a multifaceted construct, i.e., as an umbrella term that encompasses a set of fairly related but still relatively independent phenomena. This finding can also contribute to the understanding of reasons for the poor reliability of general measures of individual differences in confirmation bias (Rassin, 2008).

Besides the attempts to furtherly improve metric properties of existing measures, future studies might benefit from the inclusion of measures of confirmation bias in other processes (e.g., visual search; Rajsic et al., 2015), but also from the inclusion of other experimental paradigms (e.g., argument evaluation task; Stanovich et al., 2016) and related phenomena (e.g., myside bias; Perkins, 1989). This could help not only in answering questions about the generalizability of the proneness to the confirmation bias, but also in understanding the conceptual distinction between closely related phenomena (Mercier, 2016). As we have intended to demonstrate in the present study, the conceptual and methodological apparatus of differential psychology can be useful in this endeavor.

References

Beattie, J., & Baron, J. (1988). Confirmation and matching biases in hypothesis testing. *The Quarterly Journal of Experimental Psychology*, *40*(2), 269-297.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality*, *80*(4), 796-846.

Eagly, A. H., Chaiken, S., Chen, S., & Shaw-Barnes, K. (1999). The impact of attitudes on memory: An affair to remember. *Psychological Bulletin*, *125*(1), 64–89.

Evans, J. S. B. T. (1989). *Bias in human reasoning: Causes and consequences.* London: Erlbaum.

Hoch, S. J., & Ha, Y. W. (1986). Consumer learning advertising and the ambiguity of product experience. *Journal of Consumer Research*, *13*(2), 221-233.

Klayman, J. (1995). Varieties of confirmation bias. In J. Busemeyer, R. Hastie, & D. L. Medin (Eds.), *Decision making from a cognitive perspective* (pp. 365-418). New York: Academic Press.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*(2), 211-228.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*(11), 2098–2109.

Mercier, H. (2016). Confirmation bias – myside bias. In R. F. Pohl (Ed.), *Cognitive illusions: A hand-*

book on fallacies and biases in thinking, judgment, and memory, 2nd edition (pp. 109-124). Psychology Press.

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

Oswald, M. E., & Grosjean, S. (2004). Confirmation bias. In R. F. Pohl (Ed.), *Cognitive illusions: A handbook on fallacies and biases in thinking, judgment, and memory* (pp. 79–96). Psychology Press.

Perkins, D. N. (1989). Reasoning as it is and could be: an empirical perspective. In D. Topping, D. Crowell, & V. Kabayashi (Eds.), *Thinking across Cultures* (pp. 175–194). Hillsdale, NJ: Erlbaum.

Rajsic, J., Wilson, D. E., & Pratt, J. (2015). Confirmation bias in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(5), 1353-1364.

Rassin, E. (2008). Individual differences in the susceptibility to confirmation bias. *Netherlands Journal of Psychology*, *64*(2), 87–93.

Snyder, M., & Swann, W. B. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology*, *36*(11), 1202–1212.

Stanovich, K., & West, R. F. (1998). Cognitive ability and variation in selection task performance. *Thinking & Reasoning*, *4*(3), 193-230.

Stanovich K. E., West R. F., Toplak M. E. (2016). *The rationality quotient - Toward a test of rational thinking*. Cambridge, MA: MIT Press.

Trope, Y., & Bassok, M. (1983). Information-gathering strategies in hypothesis-testing. *Journal of Experimental Social Psychology*, *19*(6), 560-576.

Stangor, C., & Mcmillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literature. *Psychological Bulletin*, *111*(1), 42–61.

Vedejová, D., & Čavojová, V. (2022). Confirmation bias in information search, interpretation, and memory recall: Evidence from reasoning about four controversial topics. *Thinking & Reasoning*, *28*(1), 1–28.

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

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology I* (pp. 106–137). Harmondsworth: Penguin.