

Results from a Confirmatory Replication Study of Bem (2011): Precognitive Detection of Erotic Stimuli?

Eric-Jan Wagenmakers, Ruud Wetzels, Denny Borsboom, Han L. J. van der Maas, and Rogier A. Kievit
University of Amsterdam

Introduction

In 2011, Dr. Bem published an article in the *Journal of Personality and Social Psychology*, the flagship journal of social psychology, in which he claimed that people can look into the future (Bem, 2011). In his first experiment, “precognitive detection of erotic stimuli”, participants were instructed as follows: “(...) on each trial of the experiment, pictures of two curtains will appear on the screen side by side. One of them has a picture behind it; the other has a blank wall behind it. Your task is to click on the curtain that you feel has the picture behind it. The curtain will then open, permitting you to see if you selected the correct curtain.” In the experiment, the location of the pictures was random and chance performance is therefore 50%. Nevertheless, Bem’s participants scored 53.1%, significantly higher than chance; however, the effect was present only for erotic pictures, and not for neutral pictures, positive pictures, negative pictures, and romantic-but-not-erotic pictures. Bem also claimed that the psi effects are more pronounced for extraverts, and that for certain erotic pictures women show psi but men do not.

We set out to replicate Bem’s experiment in a purely confirmatory fashion. First we detailed our method, design, and planned analyses in a document that we posted online before a single participant was tested.¹ As outlined in the online document, our replication focused on Bem’s key findings; therefore, we tested only women, used only neutral and erotic pictures, and included a standard extraversion questionnaire. We also tested each participant in two contiguous sessions. Each session featured the same pictures, but presented them in a different random order.² The idea is that individual differences in psi –if these exist– lead to a positive correlation between performance on session 1 and session 2. Performance is quantified by the proportion of times that the participant chooses the curtain that hides the picture. Each session featured 60 trials, with 45 neutral pictures and 15 erotic pictures.

¹See <http://confrep.blogspot.nl/> and http://dl.dropbox.com/u/1018886/Advance_Information_on_Experiment_and_Analysis.pdf.

²The online document detailed a method of yoking picture location and picture type. Due to a miscommunication with the programmer, yoking was not properly implemented and presentation of picture location and picture type was instead just random.

A vital part of the online document concerns the *a priori* specification of our analyses. First we outlined our main analysis tool, the Bayes factor t-test:

“Data analysis proceeds by a series of Bayesian tests. For the Bayesian t-tests, the null hypothesis H_0 is always specified as the absence of a difference. Alternative hypothesis 1, H_1 , assumes that effect size is distributed as Cauchy(0,1); this is the default prior proposed by Rouder et al. (2009). Alternative hypothesis 2, H_2 , assumes that effect size is distributed as a half-normal distribution with positive mass only and the 90th percentile at an effect size of 0.5; this is the “knowledge-based prior” proposed by Bem et al. (submitted).³ We will compute the Bayes factor for H_0 vs. H_1 (BF_{01}) and for H_0 vs. H_2 (BF_{02}).”

The next six sections re-iterate the predictions from our online document and present the resulting Bayes factors. In the end, we tested 100 participants who each contributed two sessions. Because we use the Bayes factor we did not have to specify the number of participants in advance.

³Current addendum: this paper has since been published (i.e., Bem, Utts, & Johnson, 2011).

Performance: Neutral vs. Erotic, Session 1

Confirmatory test 1: Based on the data of session 1 only: Does performance for erotic pictures differ from performance for neutral pictures? To address this question we compute a paired t test (Rouder, Speckman, Sun, Morey, & Iverson, 2009) and monitor BF_{01} and BF_{02} as the data come in. Figure 1 shows the results.

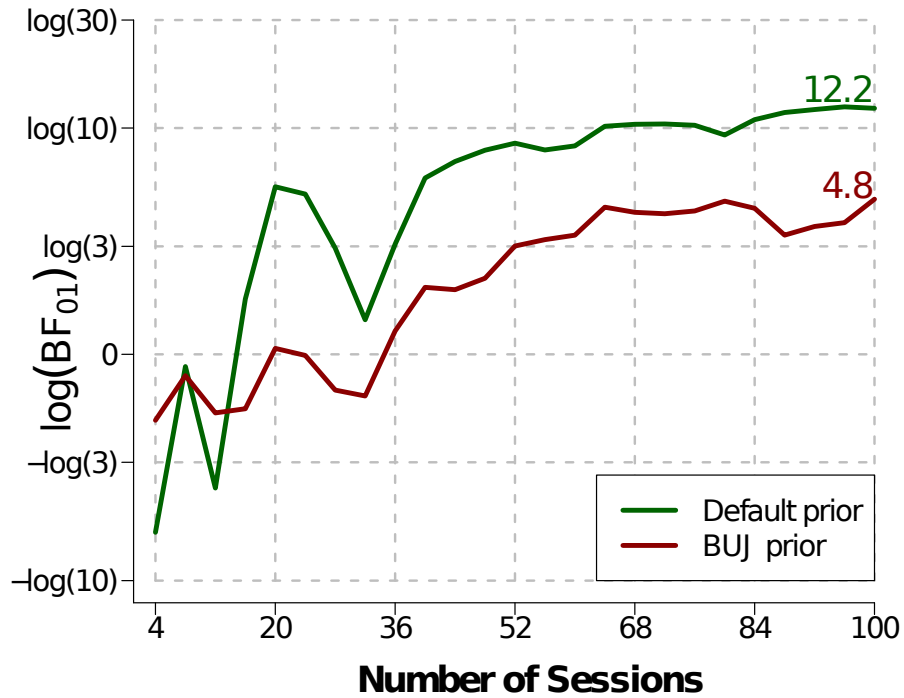


Figure 1. : Performance for erotic pictures does not differ from performance for neutral pictures (data from session 1 only). The logarithm of the Bayes factor is monitored as the data come in; $\log(BF_{01})$ is shown in green, $\log(BF_{02})$ is shown in red.

Performance: Erotic vs. Chance, Session 1

Confirmatory test 2: Based on the data of session 1 only: Does performance for erotic pictures differ from chance (in this study 50%)? To address this question we compute a one-sample t test and monitor BF_{01} and BF_{02} as the data come in. Figure 2 shows the results.

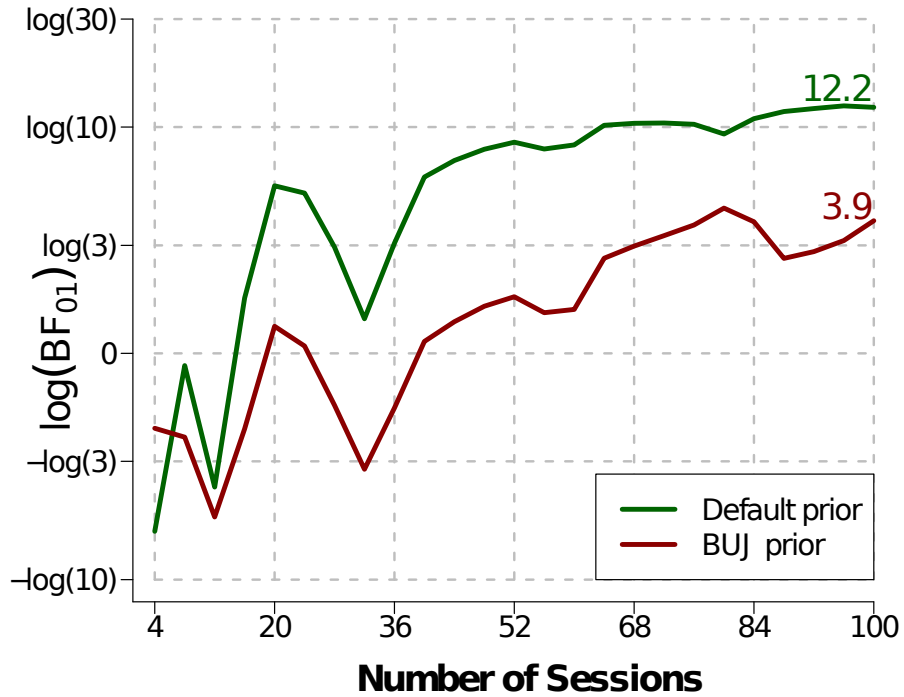


Figure 2. : Performance for erotic pictures does not differ from chance (data from session 1 only). The logarithm of the Bayes factor is monitored as the data come in; $\log(BF_{01})$ is shown in green, $\log(BF_{02})$ is shown in red.

Correlation Extraversion and Performance Erotic Pictures

Confirmatory test 3: Based on the data of session 1 only: Is there a positive correlation between extraversion scores and performance for erotic pictures? This possibility was suggested by Bem (2011), and we assess this claim using the default Bayesian test for correlation proposed by Jeffreys (1961). Figure 3 shows the results.

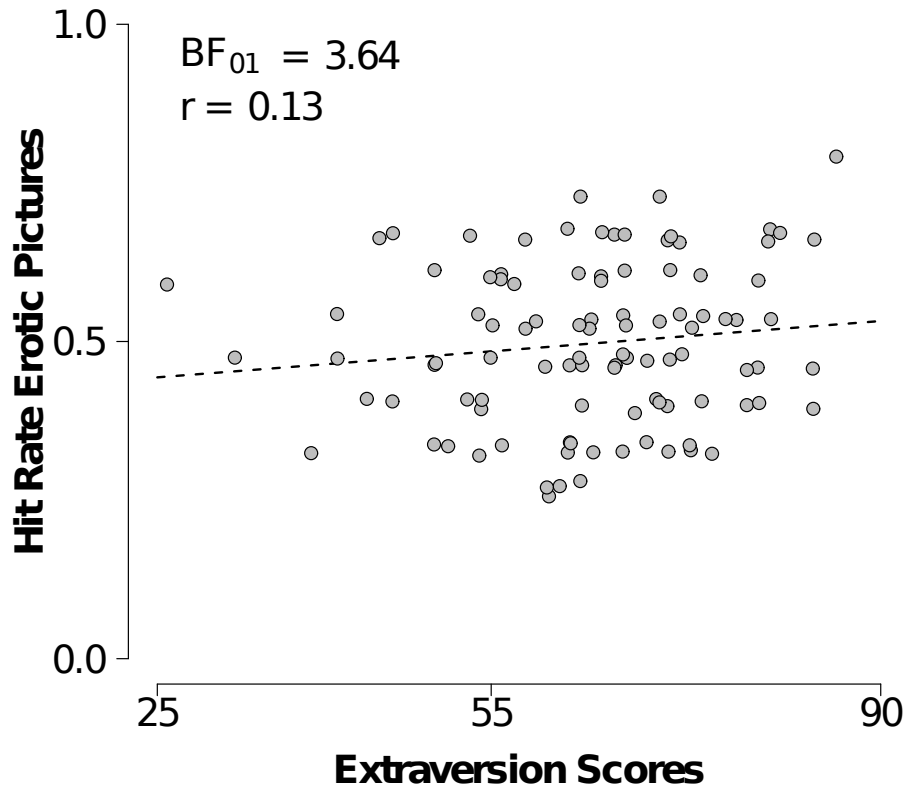


Figure 3. : There is no relation between extraversion scores and performance on erotic pictures. The correlation is 0.13, and the Bayes factor in favor of the null is 3.64. Note that this is not an order restricted test. Data are jittered to prevent visual overlap.

Correlation Between Performance Session 1 and Session 2

Confirmatory test 4: If participants have ESP, this trait should be related from session 1 to session 2. In other words, individual differences in ESP express themselves statistically as a positive correlation between performance on erotic pictures for session 1 and session 2. This prediction will again be tested using the default Bayesian test for correlation proposed by (Jeffreys, 1961). Figure 4 shows the results.

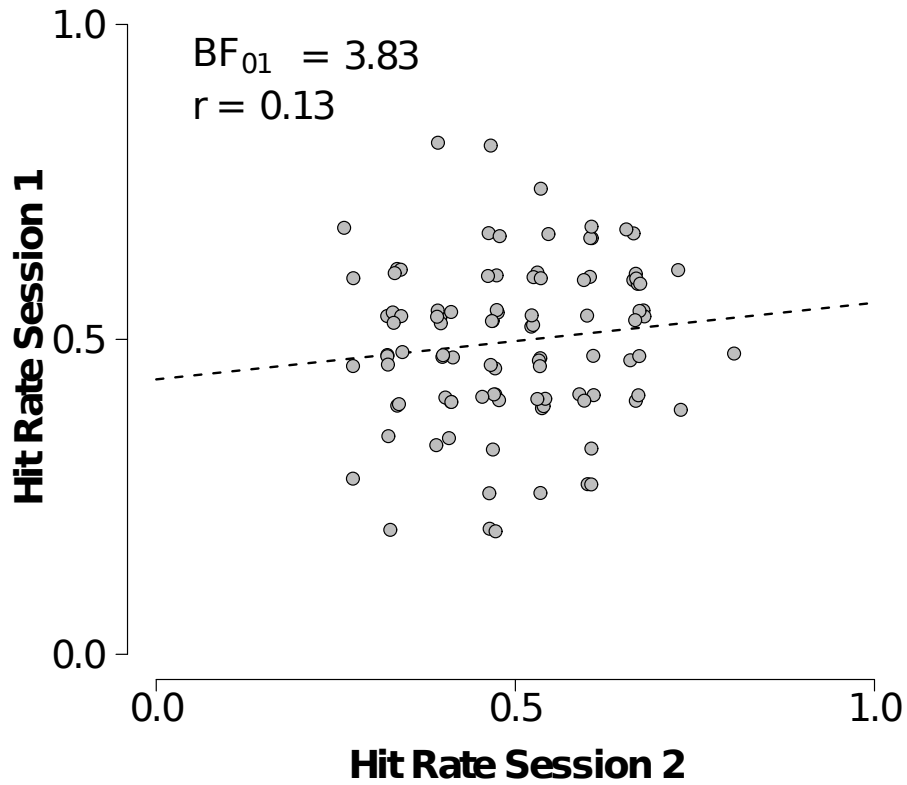


Figure 4. : There is no relation between performance on session 1 and session 2. The correlation is 0.13, and the Bayes factor in favor of the null is 3.83. Note that this is not an order restricted test. Data are jittered to prevent visual overlap.

Performance: Neutral vs. Erotic, Both Sessions Combined

Confirmatory test 5: Based on the data both sessions combined: Does performance for neutral pictures differ from performance for erotic pictures? To address this question we compute a paired t test and monitor BF_{01} and BF_{02} as the data come in. Figure 5 shows the results.

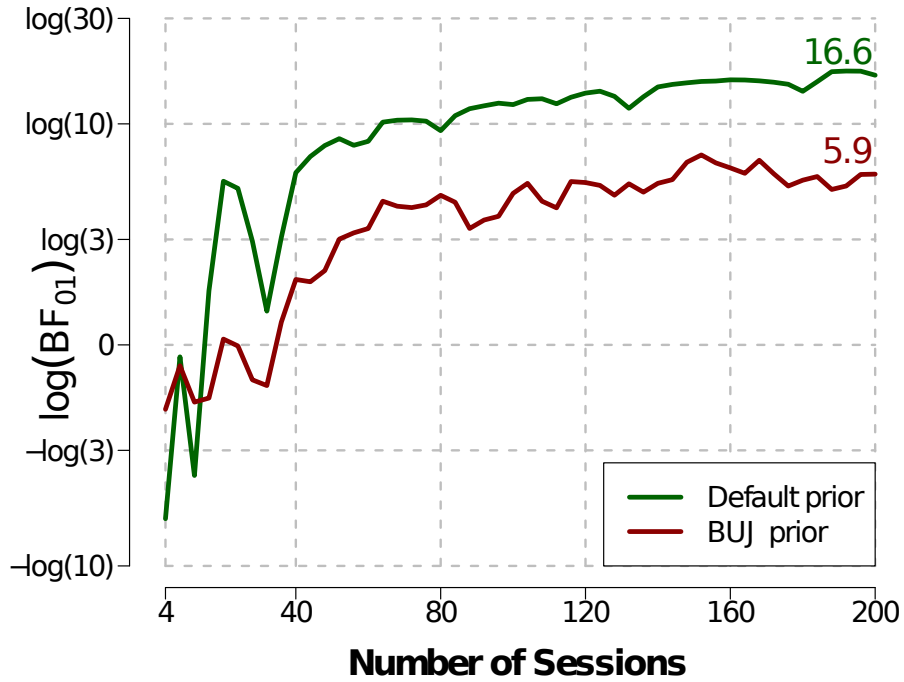


Figure 5. : Performance for erotic pictures does not differ from performance for neutral pictures (data from sessions 1 and 2). The logarithm of the Bayes factor is monitored as the data come in; $\log(BF_{01})$ is shown in green, $\log(BF_{02})$ is shown in red.

Performance: Erotic vs. Chance, Both Sessions Combined

Confirmatory test 6: Based on the data both sessions combined: Does performance for neutral pictures differ from performance for erotic pictures? To address this question we compute a paired t test and monitor BF_{01} and BF_{02} as the data come in. Figure 6 shows the results.

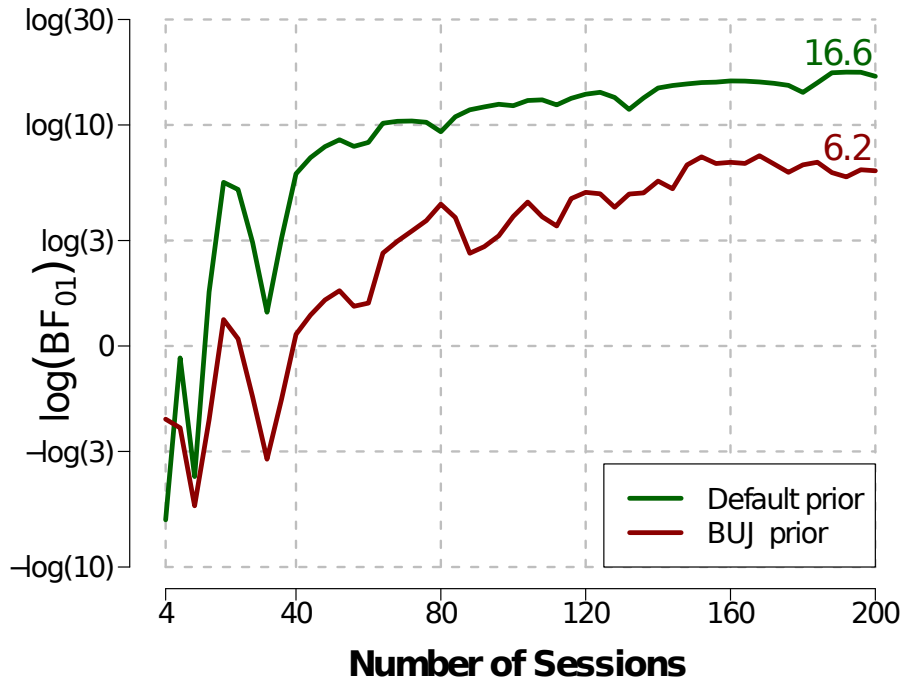


Figure 6. : Performance for erotic pictures does not differ from performance for neutral pictures (data from sessions 1 and 2). The logarithm of the Bayes factor is monitored as the data come in; $\log(BF_{01})$ is shown in green, $\log(BF_{02})$ is shown in red.

Conclusion

All tests yield evidence in favor of the null hypothesis. In other words, all confirmatory studies yielded evidence *against* the hypothesis that people can look into the future.

References

- Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*, 407–425.
- Bem, D. J., Utts, J., & Johnson, W. O. (2011). Must psychologists change the way they analyze their data? *Journal of Personality and Social Psychology*, *101*, 716–719.
- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*, 225–237.