

Biased Evaluation of Abstracts Depending on Topic and Conclusion: Further Evidence of a Confirmation Bias Within Scientific Psychology

Andreas Hergovich · Reinhard Schott ·
Christoph Burger

Published online: 29 August 2010
© Springer Science+Business Media, LLC 2010

Abstract The present paper investigated whether academic psychologists show a tendency to rate the quality and appropriateness of scientific studies more favorably when results and conclusions are consistent with their own prior beliefs (i.e., confirmation bias). In an online experiment, 711 psychologists completed a questionnaire (e.g., about their belief in astrology) and evaluated research that was presented in form of a short abstract in which 40 different behaviors (e.g., alcohol consumption, willingness to share money) have been tried to be predicted. The research to be evaluated varied on three dimensions which were all manipulated between subjects: (1) the predictors of the 40 behaviors (either Big Five or astrological factors), (2) the methodological quality of the study (low, medium, high), and (3) the results and subsequent conclusion of the study (confirmation or disconfirmation of the hypotheses). Factor-analyzed scores of participants' ratings on 8 scales, resulting in 2 factors termed quality and appropriateness, served as dependent measures. The main result of the study is a two-way interaction: Psychologists tended to evaluate results qualitatively higher when they conformed to their own prior expectations, as in this case, when astrological hypotheses were disconfirmed.

Keywords Confirmation bias · Innovation · Science · Psychologists · Astrology

The term “confirmation bias” refers to the tendency to seek out, attend to, and remember information that is in line with one's own beliefs (Oswald and Grosjean 2004). It is also known under the name conservatism (Edwards 1968), belief perseverance (Ross et al. 1975), biased assimilation (Lord et al. 1979), confirmational response bias (Epstein 2004), or agreement-effect (Koehler 1993). When confirmation bias is at work, confirmatory evidence is selectively processed and

welcomed, whereas disconfirmatory experiences are often ignored or discredited (see Nickerson 1998 for an overview of the confirmation bias in several contexts). Arguments inconsistent with prior beliefs are subjected to more skeptical analysis and consequently judged to be weaker than arguments compatible with prior beliefs (Hart et al. 2009). Thus, persons biased towards their own hypotheses do not properly consider alternative hypotheses and disregard factors that might disprove their ideas—a mechanism that systematically impedes the possibility to reject the hypothesis held to be true.

Confirmation bias is often described as a result of automatic cognitive processing. Individuals do not use deceptive strategies to fake data, but forms of information processing that take place more or less unintentionally (Oswald and Grosjean 2004). According to Maccoun (1998), most biased evidence processing occurs unintentionally through a combination of both “hot” (i.e., motivated) and “cold” (i.e., cognitive) mechanisms.

Confirmation Bias in Academia

In contrast to the common cliché of the disinterested, unprejudiced, objective and rational scientist, it has been repeatedly shown that confirmation bias is not restricted to laymen or individuals suffering from clinical disorders, but is also manifest in the realm of science itself. Both experts and novices are affected, although evidence suggests that experts in a field show reduced confirmation bias when compared to novices (Krems and Zierer 1994; Van Ophuysen 2006). Previous research has shown that the assessment of the quality of scientific studies seems to be particularly vulnerable to confirmation bias. It has been found several times that scientists rate studies that report findings consistent with their prior beliefs more favorably than studies reporting findings inconsistent with their previous beliefs (e.g., Epstein 2004; Goodstein and Brazis 1970; Greenwald et al. 1986; Koehler 1993; Mahoney 1977). However, assuming that the research question is relevant, the experimental design adequate and the data are clearly and comprehensively described, the found results should be of importance to the scientific community and should not be viewed prejudicially—regardless of whether they conform to current theoretical predictions (Mahoney 1977). One could even say that research with the potential to produce controversial findings is especially important to progress in the sciences (Armstrong 1996). But scientific innovators often meet with resistance from the scientific community (Armstrong 1996; also see Horrobin 1990 for numerous examples of harsh peer review given to research presenting controversial results). Therefore, confirmation bias may be especially harmful to objective evaluations regarding nonconforming results, since biased individuals might regard opposing evidence to be weak in principle and give little serious thought to revising those beliefs (Koehler 1993).

Previous Studies on Conformational Bias in Academia

In their seminal study, Goodstein and Brazis (1970) sent 1,000 psychologists randomly one of two virtually identical abstracts of presumably empirical research

A. Hergovich (✉) · R. Schott · C. Burger
Faculty of Psychology, University of Vienna, Liebiggasse 5, 1010 Vienna, Austria
e-mail: andreas.hergovich@univie.ac.at

on astrology. Whereas one abstract reported significant positive correlations between planetary configurations at the time of birth and subsequent choice of vocation concluding that additional research in this area would be fruitful, the other abstract reported no significant relationships, concluding that future research was unlikely to be productive. The psychologists had to rate the abstracts. The authors hypothesized that psychologists would be more willing to accept findings disconfirming the validity of astrology than findings confirming it. Results showed that subjects receiving the abstract with disconfirming findings rated the study as significantly better designed, as having more validity, and as containing more adequate conclusions than subjects receiving the abstract with confirming findings.

Mahoney (1977) conducted a study, where 67 reviewers assessed different versions of a report on a fictitious experiment. He found that reviewers were strongly biased against the papers that reported results contrary to their common accepted theoretical perspective (i.e., behaviorist). They rejected them on the basis of poor methodology while accepting papers with confirmatory outcomes that used the identical methodology. The authors conclude that the assessment of scientific research appears to be dependent on the correspondence between a scientist's belief and the results of the study and that this bias may be one of the most pernicious and counterproductive elements in the social sciences.

Moss and Butler (1978) conducted a survey measuring belief in a paranormal phenomenon (i.e., extrasensory perceptive powers, ESP). They found that academic psychologists were much more skeptical regarding paranormal phenomena than other groups. In the study of Koehler (1993), scientists judged studies that were in line with their prior beliefs (i.e., disconfirmation of parapsychological theories) to be more relevant, methodologically sound and clearly presented than otherwise identical studies that were not in line.

The Present Study

The current study aims to examine whether academic psychologists of today would be similarly biased as their counterparts in previous studies (e.g., Goodstein and Brazis 1970; Koehler 1993; Moss and Butler 1978) and is a conceptual replication of Goodstein and Brazis' (1970) seminal study, providing some enhancements to overcome methodological limitations and rule out alternative explanations. In the above mentioned studies, it has been shown several times that a study disconfirming a controversial, non-established theory is rated more favorably by psychologists than the same study confirming it. Prior studies concluded that this would be a manifestation of confirmation bias. But what is missing so far is evidence that the disconfirmation of a non-established theory is rated by psychologists more favorably than the disconfirmation of an established theory as well as that the confirmation of an established theory is rated more favorably than the confirmation of a non-established theory, respectively. Another problem could be that different factors than the non-agreement with the non-established theory led the scientists to rate the study with the disconfirming results more favorably. Such factors could, for example, be found in the methodology of the to-be-assessed study, such as the study design, statistical analyses or number of participants, which could make disconfirming results more plausible.

The novel aspect of our study is therefore the implementation of a new control group: In addition to the confirmation or disconfirmation of a less accepted theory (i.e., astrology), we also used abstracts confirming or disconfirming a widely accepted psychological personality theory (i.e., Big Five). A further important advancement was the variation of the methodological quality of the abstracts on three levels (i.e., small, medium and large sample; different statistical methods) between subjects in order to be able to discern whether potential differences in evaluations may be simply artifacts of abstract quality or methodological considerations.

Hypotheses

In the present study, the following hypotheses were proposed:

- H1: In line with the findings of Goodstein and Brazis (1970) and Koehler (1993) and in correspondence with Moss and Butler's (1978) claim that skepticism is most pronounced among academic psychologists, we expected that an abstract confirming the theory of astrology will be viewed much more critically than one confirming a prominent personality theory. More precisely we hypothesized an interaction between the topic of study (i.e., hypotheses pertaining to astrology or personality theory) and its conclusion (i.e., whether the underlying hypotheses will be confirmed or disconfirmed by the results reported in the abstract). Abstracts supporting the theory of astrology should be given less favorable ratings, whereas abstracts confirming personality theory should be assessed more favorably.
- H2: In correspondence to the findings of Krems and Zierer (1994) and Van Ophuysen (2006) that experts are less influenced by the confirmation bias than novices, we hypothesized that psychologists on a higher level of their academic career should be less influenced by their prior beliefs than their colleagues ranking lower on the academic hierarchy (i.e., a 3-way interaction between *topic*, *conclusion* and *professional education*).
- H3: In order to rule out that potential differences in evaluations may be simply artifacts of abstract quality or methodological considerations, there should neither be any interaction between the *methodological quality of the abstract* and the *conclusion* (i.e., confirming and disconfirming) nor between the *methodological quality of the abstract* and the *subject* of the abstracts (i.e., astrology and personality theory). The three different levels of quality of abstract should, however, influence the assessment of the abstracts. The higher the quality of the abstracts, the better they should be evaluated (i.e., main effect of the variable "quality of abstracts").

Method

Participants

The study was undertaken as an online investigation. On the whole, 10,056 email-addresses of potential participants were obtained through publicly available

information on websites of 46 universities in the USA, Canada, Australia, New Zealand, Great Britain, Ireland, Switzerland and Germany. Subjects were contacted personally via e-mail and were invited to participate in the study. Subjects entered the online questionnaire via a link and were randomly assigned to one of the experimental conditions. The following criteria had to be met to be included in the study: Subjects had to (1) possess at least a diploma in psychology and (2) be employed by a psychological research institute of a university.

A total of 1,060 (response rate 10.5%) of the contacted individuals followed the hyperlink and began to participate in the online-study. 724 subjects (completion rate 68.3%) finished the study by pressing the send-button at the last page of the questionnaire. Thirteen subjects had to be excluded from the sample because they did not possess a diploma in psychology or because it was their own wish.

Eventually, 711 (45.3% female; 54.7% male) subjects were included in the study. The age of the participants ranged from 22 to 85 years, with a mean age of 44.44 years ($SD=13.03$). Regarding the academic level, 113 (15.8%) were graduates, 213 (30.0%) doctors (PHD), 103 (14.5%) Assistant Professors, 100 (14.1%) Associate Professors and 182 (25.6%) Professors.

Materials

In the online-experiment, participants had to evaluate a short research abstract in which 40 different behaviors, like for example alcohol consumption, have been tried to be predicted. The abstract varied on three dimensions which were all manipulated between subjects, that is, (1) the predictors of the 40 behaviors (either Big Five or astrological factors), (2) the methodological quality of the study (low, medium, high), and (3) the conclusion (confirmation or disconfirmation of the hypotheses). The low-quality version reported a study with only 27 subjects and method of analysis was only correlation. The medium-quality version included 47 subjects, more complex and more detailed reports and the method of analysis was regression-analysis. The high-quality version included 197 subjects and employed more detailed and complex analyses. Thus a 3 (methodological quality: low, medium, high) \times 2 (topic: astrology, personality theory) \times 2 (conclusion: confirmation, disconfirmation) design with 12 abstracts resulted. All abstracts were based on an article by Paunonen and Ashton (2001) and are attached in the Appendix. The distributions of the subjects to the conditions of the study can be seen in Table 1.

Procedure and Measurements

At the beginning, sociodemographic information of the participants (i.e., sex, age, academic degree, area of work, level of professional experience) was solicited. As a next step, their belief in astrology was assessed using the following items: “Do you believe in astrology” (ranging from “not at all” [1] to “very strongly” [5]), “Do you believe that astrology can predict the future?” (ranging from “not at all” [1] to “very strongly” [5]), “Do you believe that the constellation of the stars at the time of birth has an influence on the personality of a person” (ranging from “not at all” [1] to “very strongly” [5]), and “Do you read horoscopes in newspapers, magazines or the internet?” (ranging from “not at all” [1] to “very strongly” [5]).

Table 1 Distribution of the participants to the different conditions of the study

	Low quality		Medium quality		High quality		Total
	Confirm	Disconfirm	Confirm	Disconfirm	Confirm	Disconfirm	
Astrology	55	79	61	61	61	52	366
Big Five	70	64	51	54	52	54	345
Total	125	143	109	115	113	106	711

“Confirm” = hypotheses in the abstract were confirmed; “Disconfirm” = hypotheses in the abstract were disconfirmed

After reading the abstract, subjects were asked to rate the described study on the following dimensions using five-point rating scales ranging from “very low” (1) to “very high” (5) (also see Goodstein and Brazis 1970): (a) quality of experimental design, (b) validity of results, (c) appropriateness of conclusions, (d) significance for future research, (e) believability of the results, (f) appropriateness for publication in psychological journals, (g) appropriateness of topic for psychological research, (h) appropriateness for own research. Eventually, the subjects were offered the possibility to make an open comment about the study.

Results

Subjects receiving an abstract confirming the theory of astrology abandoned the questionnaire more often (attrition rate: 38.9%) than subjects receiving an abstract refuting the theory of astrology (attrition rate: 27.5%) (see Table 2, $Chi^2=8.02$, $p<.05$). For the subjects receiving the abstract on personality theory, no significant differences in drop-out rates emerged.

First, all items measuring belief in astrology were averaged to yield a 5-point index of belief in astrology. A descriptive analysis showed that the vast majority of the participants did not believe in astrology—only 14 participants (2.0%) showed a value above the midpoint. However, since the integrity of our manipulations depended on participants not believing in astrology, data for participants who rated their belief in astrology above the midpoint were omitted for the following analyses.

Second, all items pertaining to the assessment of abstracts were subjected to a factor analysis. The two resulting factors together explained 64.60% of the

Table 2 Drop-out rates for the subjects receiving the abstracts pertaining to astrology

	Drop-out rate	Participation rate	Total
Confirmation	111 (38.9%)	174 (61.1%)	285
Disconfirmation	73 (27.5%)	192 (72.5%)	265
Total	184	366	550

Values in brackets represent the percent of participants relative to the total number of participants in the confirmation or disconfirmation condition

variance. The first factor (explained variance: 45.03%) highly loaded on the items measuring “quality of design,” “validity of results,” “appropriateness of the conclusions,” “significance of the study for future research,” “believability of the results” and “appropriateness for publication in a psychological journal” and was therefore labeled “perceived quality of abstract.” The second factor (explained variance: 19.57%) was characterized by high loadings on the variables “appropriateness for research by psychologists in general” and “appropriateness for own research” and was labeled “perceived appropriateness of abstract.” Internal consistencies (Cronbach’s Alpha) were .85 for the factor “perceived quality of abstract” and .57 for the factor “perceived appropriateness of abstract.” The answers of the items pertaining to the two factors were averaged to yield an index of “perceived quality of abstract” and “perceived appropriateness of abstract.”

Statistical Analyses

For the sake of brevity and clarity, the following analyses will focus only on those effects that achieved a significance level of 0.05 or less. In order to test the hypotheses, we conducted a 2 (*topic*: astrology, personality theory) \times 3 (*methodological quality of abstract*: high, medium, poor) \times 2 (*conclusion*: confirmation, disconfirmation) \times 2 (*sex*: male, female) \times 3 (*professional education*: Graduate, PHD and Assistant Professor, Professor and Associate Professor) multivariate analysis of variance (MANOVA) with the dependent variables “perceived quality of abstract” and “perceived appropriateness of abstract.” This analysis revealed significant multivariate effects of *topic* ($F(2, 624)=51.38, p<.05; \eta^2=.141$), *methodological quality of abstract* ($F(4, 1,250)=3.39, p<.05; \eta^2=.011$), *conclusion* ($F(2, 624)=6.40, p<.05; \eta^2=.020$), *professional education* ($F(4, 1,250)=4.22, p<.05; \eta^2=.013$), and a two-way interaction between *topic* and *conclusion* ($F(2, 624)=19.15, p<.05; \eta^2=.058$). *Sex* had no significant effect ($F(2, 624)=1.13, p>.05; \eta^2=.004$). Because of the significant interactions the highly significant main-effects of *topic* and *conclusion* will not be interpreted.

Univariate Effects for Abstract Quality

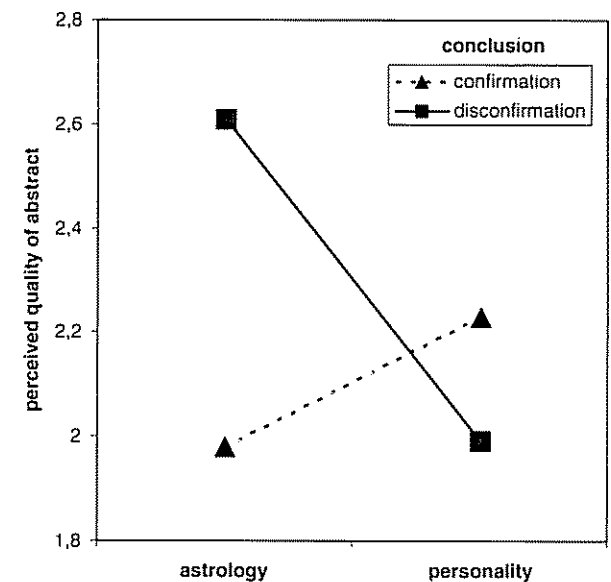
For the dependent variable “perceived quality of abstract” main effects of topic, methodological quality of abstract, conclusion and professional education as well as a highly significant interaction between topic and conclusion could be found at the univariate level. The different levels of methodological quality of abstract (low, medium, high) had a significant effect on the assessed “perceived quality of abstract” ($F(2, 625)=6.81, p<.05; \eta^2=.021$). Post-Hoc-Tests (Scheffé) revealed that both the abstracts of medium ($M=2.25, SEM=0.07$) and high methodological quality ($M=2.33, SEM=0.07$) were assessed significantly better than the abstract of low quality ($M=2.04, SEM=0.05$). Furthermore, professional education (Graduate, PHD and Assistant Professor, Professor and Associate Professor) had a significant effect on the assessed “perceived quality of abstract” ($F(2, 625)=7.24, p<.05; \eta^2=.023$). Post-Hoc-Tests (Scheffé) revealed that graduates ($M=2.35, SEM=0.08$) and PhDs and assistants ($M=2.23, SEM=0.04$) assessed the abstracts significantly better than professors did ($M=2.03, SEM=0.05$).

The interaction between topic and conclusion was highly significant ($F(1, 625)=37.05, p<.05; \eta^2=.056$). Simple effect tests showed that all possible comparisons were significant. As shown in Fig. 1, the quality of the abstract disconfirming the theory of astrology (consistent; $M=2.61, SEM=0.08$) was assessed much higher than the quality of the abstract disconfirming the personality theory (inconsistent; $M=1.99, SEM=0.07; F(1,705)=55.40, p<.05$). For the abstracts with confirming conclusions, a reversed pattern emerged: The abstract confirming personality theory was assessed higher (consistent; $M=2.23, SEM=0.08$) than the abstract confirming the theory of astrology (inconsistent; $M=1.98, SEM=0.07; F(1,705)=11.46, p<.05$). Furthermore, abstracts about the theory of Big Five were rated more favorably when confirmed than when not confirmed ($F(1,705)=15.13, p<.05$), whereas abstracts about the theory of astrology were rated less favorable when confirmed than when not confirmed ($F(1,705)=48.28, p<.05$). Partial Eta-Square as a measure of effect-size for this interaction was .056 (effect size $f=.24$) indicating a moderate effect (Cohen 1988).

Univariate Effects for Abstract Appropriateness

For the dependent variable “perceived appropriateness of abstract” main effects of *topic*, *conclusion*, *professional education* and, although not significant at the multivariate level, the interaction between *conclusion* and *professional education* was significant at univariate level. The different levels of professional education (Graduate, PHD and Assistant Professor, Professor and Associate Professor) had a significant effect on the assessed “perceived appropriateness of abstract” ($F(2, 625)=4.32, p<.05; \eta^2=.014$). Post-Hoc-Tests (Scheffé) revealed that graduates ($M=2.37, SEM=0.10$) assessed the abstracts significantly better than professors did ($M=2.14, SEM=0.06$). PHD and Assistants’ judgments ($M=2.35, SEM=0.05$) did not differ from those of graduates and professors.

Fig. 1 Interaction between topic (astrology theory, personality theory) and conclusion (confirmation, disconfirmation) on perceived quality of the abstracts



The interaction between conclusion and professional education ($F(2, 625)=5.68$, $p<.05$; $\eta^2=.018$; see Fig. 2) was also significant. A simple effect test showed that Graduates rated abstracts confirming hypotheses—regardless if they pertained to astrology or personality theory—as less appropriate ($M=2.04$, $SEM=0.13$) than abstracts refuting their hypotheses ($M=2.70$, $SEM=0.14$; $F(1,705)=6.29$, $p<.05$). PhDs and Assistant Professors assessed the appropriateness of abstracts confirming their hypotheses ($M=2.33$, $SEM=0.07$) very similar to those refuting their hypotheses ($M=2.37$, $SEM=0.07$; $F(1,705)=0.04$, $p>.05$), just like Professors and Associate Professors who rated abstracts with positive results to be similarly appropriate ($M=2.19$, $SEM=0.087$; $F(1,705)=0.83$, $p>.05$) as abstracts with negative results ($M=2.09$, $SEM=0.07$).

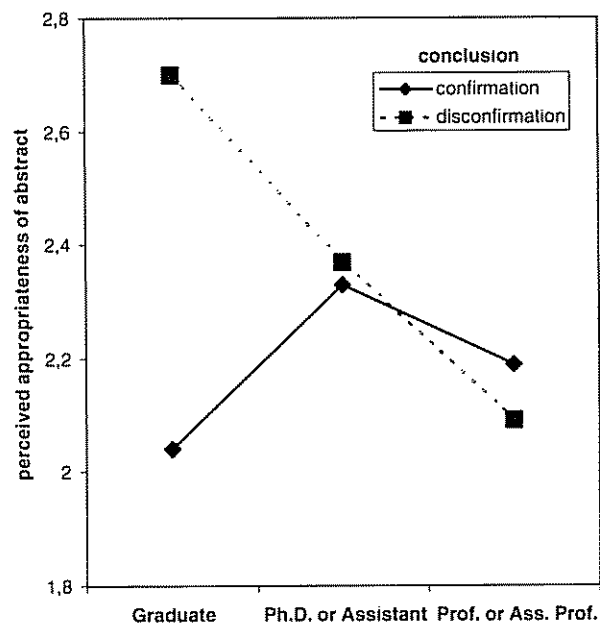
Discussion

The aim of the present study was to examine whether academic psychologists of today would show a similar tendency to rate the quality and appropriateness of scientific studies more favorably if results and conclusions were consistent with their own prior beliefs (i.e., confirmation bias) as their counterparts in previous studies (e.g., Goodstein and Brazis 1970; Koehler 1993; Moss and Butler 1978).

Discussion of Hypotheses

Hypothesis 1 Concerning the main thesis of the paper, support was found as the abstracts confirming a disputed theory (i.e., theory of astrology) were assessed less

Fig. 2 Interaction between professional education (Graduates, PHD and Assistants, Professors and Associate Professors) and conclusion (confirmation, disconfirmation) on perceived appropriateness of the abstracts



favorably in terms of quality than the same abstracts confirming personality theory. By the same token, abstracts disconfirming astrological hypotheses were assessed more favorably than abstracts disconfirming hypotheses based on personality theory. Since all participants did not believe in the theory of astrology, we assume that the notion that astrological hypotheses were able to predict several behavioral indices from astrological factors was rather surprising. However, this cannot be accepted as a scientific justification for criticizing the quality of an abstract which is described in exactly the same way as another abstract but which on the other hand is evaluated much better. Therefore, this result can be interpreted as manifestation of the confirmation-bias. A possible explanation could be that the academic psychologists in our sample subjected the abstracts inconsistent with their prior beliefs to more skeptical analysis (Hart et al. 2009). Due to the brevity of the abstracts, a lot of details and standard operations which are vital to the studies' quality cannot be reported. In the case of consistent conclusions, participants might have interpreted the missing information in favor of the quality of the study whereas otherwise they might have not.

Moreover, no interaction between topic and conclusion could be found regarding the perceived appropriateness for psychological research and for own research. This is particular interesting, because in contrast to the abstract quality (which should be rated objectively and independently from the conclusion), the ratings on this dimension would legitimately allow the rater to incorporate his subjective feelings towards astrology. These findings could be explained by the fact that confirmation bias occurred unintentionally or that participants did not want to admit to themselves that their rating of abstract quality was influenced by their attitude towards astrology.

Hypothesis 2 The hypothesis that psychologists higher on the academic hierarchy are less influenced by the confirmation bias could not be confirmed since the expected 3-way interaction between topic, conclusion and professional education was non-significant. However, we found that scientists higher on the academic hierarchy (i.e., professors) generally assessed the quality of abstracts more critically than graduates and PhDs irrespective of the topic and whether the theory was confirmed or disconfirmed. In terms of the evaluation of the appropriateness of the abstract, we found a significant interaction between conclusion and professional education. Whereas graduates rated abstracts with positive results—regardless if the hypotheses pertained to astrology or personality theory—as less appropriate than abstracts with negative results, PhDs, assistants and professors did not make this distinction.

Hypothesis 3 The found confirmation bias could not be explained alternatively as artifacts of abstract quality or methodological considerations, because we neither found a significant interaction between abstract quality and conclusion nor between abstract quality and subject on any of the two dependent variables. Additionally, we found a significant main effect of methodological quality of abstract on the perceived abstract quality, which was also in line with H3, since we expected the participants to rate the different levels according to their quality.

Overall Discussion

The findings of the present study have four major implications. First, the results clearly confirmed that academic psychologists of today are similarly biased as their counterparts in previous studies (e.g., Goodstein and Brazis 1970; Koehler 1993). Academic psychologists rated studies that report findings consistent with their prior beliefs more favorably than studies reporting findings inconsistent with their previous beliefs. This is especially true for research with the potential to produce controversial findings, which on the other hand is vital for the progress in psychological science (Armstrong 1996). In the worst case this could lead to the suppression of new information over a long period, because of often harsh peer review given to research presenting controversial results (Horrobin 1990).

Second, we found that professors, assistants, PhDs and graduates were similarly affected by the confirmation bias. However, professors generally evaluated the abstracts less favorably in terms of quality. In terms of appropriateness, PhDs, Assistants and Professors were less influenced by the conclusion (i.e., confirmation or disconfirmation) than graduates.

Third, we could rule out alternative explanations for the found confirmation bias in the previous studies. We found evidence that academic psychologists rated the disconfirmation of a non-established theory more favorably than the disconfirmation of an established theory as well as that they rated the confirmation of an established theory more favorably than the confirmation of a non-established theory, respectively. Furthermore we could show that methodological factors were not responsible for the found results.

Finally, the findings of the present study stress that confirmation bias can have systematic detrimental effects on the development of scientific psychology. Therefore it is of utmost importance to take measures against it. A possible measure would be to encourage scientific journals to implement alternative modes of the review process, such as the *early acceptance procedure* as described by Armstrong (1996) which is based on the reviewing of the design of a study and can be done before the study is conducted (for a more detailed description of the early acceptance procedure, see Armstrong 1996). Future studies should aim to find further starting-points and methods to reduce the influence of confirmation bias within the scientific review process.

Limitations

As with other empirical studies, there are limitations that need to be acknowledged and addressed regarding the present study. The first limitation concerns the low response rate of 10.5%. A connected limitation is the fact that our sample consists of self-selected academic psychologists and does therefore not permit a straightforward generalization to the overall population of academic psychologists. Yet, we believe that this is a general vulnerability of online data collection, not something specific to the present study. Apart from this weakness, the chosen approach does also have advantages. It would not otherwise have been possible to reach a considerable big sample (>700) of academic psychologists working in different countries. We think that such a study can only be feasibly accomplished in this medium.

Since the invitation email did not explicitly mention the topic of the questionnaire, it can be safely assumed that interest or disinterest in the particular topic of the study did not directly influence the number of people starting the survey. This fact, did however, quite obviously, influence the rate of people abandoning the survey. A good indicator for this may be that subjects receiving an abstract confirming the theory of astrology abandoned the questionnaire significantly more often than subjects receiving an abstract refuting theory of astrology (see Table 2). This observation also speaks to the validity of our hypothesis that psychologists prefer the disconfirmation to the confirmation of new results that do not fit to the established prior belief. Furthermore, more positively, self-selection can be expected to render our study a strong theory test because it leads to larger sample homogeneity, which on the other hand attenuates statistical power. Therefore it is very likely that our results have underestimated the true results.

Finally, although we provided the participants with the opportunity to comment on the present study, we did not explicitly ask the participants what factors were relevant for their rating decisions. Additional questions aiming at the reasons why participants rated the quality and appropriateness of a certain abstract more or less favorably could bring more insight into the underlying mechanisms.

Appendix

List of Abstracts

Abstract 1: Astrology, Disconfirmation, High Quality

The purpose of this study is to evaluate the relative contributions of astrological determinants and astrological facets to the prediction of behavior. The authors compared five major astrological factors (sun sign, Ascendent, Medium Coeli, planets and major aspects) with five unfamiliar facets of astrology (minor aspects, lunar nodes, Asteroids, Uranian planets, Arabic parts) that constitute the astrological horoscope on their ability to predict 40 behavior criteria (e.g., alcohol consumption, willingness to share money). Both the major factors and the unfamiliar facets (the number of unfamiliar facet predictors was limited to the number of the major predictors) predicted no substantial numbers of criteria.

Method: A total of 187 subjects (97 female and 90 male university students with a mean age of 20.8) completed a questionnaire indicating their birth date, time and location of birth and a behavioral report form. As mentioned, this study compares the major astrological factors with unfamiliar facets of astrology in terms of their criterion predictability. For each subject a horoscope was calculated and 3 astrologers rated them separately for any astrological factor and facet in the 40 behavior criteria.

Results (excerpts): Because the interrater-reliability between the 3 astrologers was .83, the data of the three astrologers were collapsed together. In the first regression analyses all five major factor scales to predict each of the 40 behavior criteria were used. These five major predictors were able to account for significant amounts of variance in only 2 of the 40 criterion variables. These behavior criteria were: Parties

attended, $R1=.29$; $R0=.00$, $\%=6.4$, $p<.05$; and dieting behavior, $R1=.24$; $R0=.11$, $\%=6.2$, $p<.05$. ($R1$ = multiple correlation of sex and the predictors with criterion; $R0$ = multiple correlation of sex alone with the criterion; $\%$ = percentage of increase in criterion variance accounted for by $R1$ over $R0$.)

In the second regression analyses the five unfamiliar facets of the astrological horoscope were used to predict each of the 40 behavior criteria. The five facets were able to account for significant amount of variance in just 1 of the 40 criterion variables. This behavior criterion was: Fraternity interest, $R1=.22$, $R0=.04$, $\%=5.8$, $p<.05$. ($R1$ = multiple correlation of sex and the predictors with criterion; $R0$ = multiple correlation of sex alone with the criterion; $\%$ = percentage of increase in criterion variance accounted for by $R1$ over $R0$.)

The five major astrological factors do not account for a substantial amount of variation in human behavior. Astrological interpretations of these factors, therefore, may be useful tools for predicting and understanding behavior, but the few previous positive findings in this area are not supported in this study. There is also no reliable astrological based variation in behavior that can be traced back to the unfamiliar aspects of astrology. Additional research in this area does not appear to be fruitful.

Abstract 2: Astrology, Disconfirmation, Medium Quality

The purpose of this study is to evaluate the relative contributions of astrological determinants and astrological facets to the prediction of behavior. The authors compared five major astrological factors (sun sign, Ascendent, Medium Coeli, planets and major aspects) with five unfamiliar facets of astrology (minor aspects, lunar nodes, Asteroids, Uranian planets, Arabic parts) that constitute the astrological horoscope on their ability to predict 40 behavior criteria (e.g., alcohol consumption, willingness to share money). Both the major factors and the unfamiliar facets (the number of unfamiliar facet predictors was limited to the number of the major predictors) predicted no substantial numbers of criteria.

Method: A total of 47 undergraduate university students completed a questionnaire indicating their birth date, time and location of birth and a behavior report form. As mentioned, this study compares the major astrological factors with five unfamiliar facets of astrology in terms of their criterion predictability. An astrologer was asked to rate the expected association between the diverse unfamiliar astrological facets and the criteria, and then chose the highest rated facets as predictors of each criterion. For each subject a horoscope was calculated and an astrologer rated them separately for any astrological factor and facet in the 40 behavior criteria.

Results: In the first regression analyses all five major factor scales to predict each of the 40 behavior criteria were used. These five major predictors were able to account for significant amounts of variance only in 2 of the 40 criterion variables. These behavior criteria were: Parties attended, $R=.29$; $p<.05$; and dieting behavior, $R=.24$; $p<.05$. (R = multiple correlation of the predictors with criterion.)

In the second regression analyses the five unfamiliar facets of the astrological horoscope were used, rated by the astrologer, to predict each of the 40 behavior criteria. The five facets were able to account for significant amounts of variance in just 1 of the 40 criterion variables. This behavior criterion was: Fraternity interest, $R=.22$, $p<.05$. (R = multiple correlation of the predictors with criterion.)

The five major astrological factors do not account for a substantial amount of variation in human behavior. Astrological interpretations of these factors, therefore, may be useful tools for predicting and understanding behavior, but the few previous positive findings in this area are not supported in this study. There is also no reliable astrological based variation in behavior that can be traced back to the unfamiliar aspects of astrology. Additional research in this area does not appear to be fruitful.

Abstract 3: Astrology, Disconfirmation, Low Quality

The purpose of this study is to evaluate the relative contributions of astrological determinants to the prediction of behavior. The authors examined five coefficients of astrological factors (sun sign, Ascendent, Medium Coeli, planets and major aspects) on their ability to predict 40 behavior criteria (e.g., alcohol consumption, willingness to share money). The astrological personality factors predicted no substantial numbers of criteria.

Method: A total of 27 undergraduate university students completed a questionnaire indicating their birth date, time and location of birth and a behavior report form. For each subject a Horoscope was calculated and an astrologer rated them separately for any astrological factor in the 40 behavior criteria.

Results: Correlations were computed with each coefficient of astrological factors to predict each of the 40 behavior criteria. These astrological predictors were able to account for significant amounts of variance in only 2 of the 40 criterion variables. These behavior criteria were: Parties attended with sun sign, $r=.29$; $p<.05$; and dieting behavior with major aspects, $r=.24$; $p<.05$.

The five major astrological factors do not account for a substantial amount of variation in human behavior. Astrological interpretations of these factors are no useful tools for predicting and understanding behavior. Additional research in this area does not appear to be fruitful.

Abstract 4: Astrology, Confirmation, High Quality

The purpose of this study is to evaluate the relative contributions of astrological determinants and astrological facets to the prediction of behavior. The authors compared five major astrological factors (sun sign, Ascendent, Medium Coeli, planets and major aspects) with five unfamiliar facets of astrology (minor aspects, lunar nodes, Asteroids, Uranian planets, Arabic parts) that constitute the astrological horoscope on their ability to predict 40 behavior criteria (e.g., alcohol consumption, willingness to share money). Both the major factors and the unfamiliar facets predicted substantial numbers of criteria, but the latter did noticeably better in that regard. Moreover, the criterion variance accounted for by the unfamiliar astrological facets often included large portions not predicted by the major astrological factors. The unfamiliar facets, therefore, were able to substantially increase the maximum prediction achieved by the major factors. The results of this study are interpreted as supporting a more detailed approach to astrological assessment, one that goes beyond the interpretation of the major factors alone.

Method: A total of 187 subjects (97 female and 90 male university students with a mean age of 20.8) completed a questionnaire indicating their birth date, time and

location of birth and a behavioral report form. As mentioned, this study compares the major astrological factors with unfamiliar facets of astrology in terms of their criterion predictability. For each subject a horoscope was calculated and 3 astrologers rated them separately for any astrological factor and facet in the 40 behavior criteria.

Results (excerpts): Because the interrater-reliability between the 3 astrologers was .83, the data of the three astrologers were collapsed together. In the first regression analyses all five major factor scales to predict each of the 40 behavior criteria were used. These five major predictors were able to account for significant amounts of variance in 15 of the 40 criterion variables. These behavior criteria were for example: Parties attended, $R1=.41$, $R0=.00$, $\%=18.4$, $p<.001$; Dieting behavior, $R1=.38$, $R0=.13$, $\%=14.1$, $p<.001$; Fraternity interest, $R1=.31$, $R0=.11$, $\%=10.5$, $p<.01$; Participation in sports, $R1=.27$, $R0=.09$, $\%=8.3$, $p<.05$; Alcohol consumption, $R1=.27$, $R0=.06$, $\%=7.9$, $p<.05$; Willingness to share money, $R1=.25$, $R0=.11$, $\%=7.4$, $p<.05$; Playing musical instruments, $R1=.23$, $R0=.00$, $\%=7.1$. ($R1$ = multiple correlation of sex and the predictors with criterion; $R0$ = multiple correlation of sex alone with the criterion; $\%$ = percentage of increase in criterion variance accounted for by $R1$ over $R0$.)

In the second regression analyses the five unfamiliar facets of the astrological horoscope were used to predict each of the 40 behavior criteria. The five facets were able to account for significant amounts of variance in 17 of the 40 criterion variables. These behavior criteria were for example: Willingness to share money, $R1=.48$, $R0=.16$, $\%=20.1$, $p<.001$; Tobacco consumption, $R1=.46$, $R0=.00$, $\%=19.8$, $p<.001$; Religiosity rating, $R1=.42$, $R0=.11$, $\%=17.6$, $p<.001$; Alcohol consumption, $R1=.34$, $R0=.03$, $\%=15.3$, $p<.001$; Parties attended, $R1=.31$, $R0=.00$, $\%=12.4$, $p<.01$; Dating variety, $R1=.28$, $R0=.09$, $\%=8.9$, $p<.05$; Fraternity interest, $R1=.26$, $R0=.04$, $\%=8.4$, $p<.05$. ($R1$ = multiple correlation of sex and the predictors with criterion; $R0$ = multiple correlation of sex alone with the criterion; $\%$ = percentage of increase in criterion variance accounted for by $R1$ over $R0$.)

The five major astrological factors account for a substantial amount of variation in human behavior. Astrological interpretations of these factors, therefore, can be useful tools for predicting and understanding behavior. But there is also reliable astrological based variation in behavior that is not accounted for by the major factors. This additional variation is evident in the specific variance residing. Additional research in this area appears to be fruitful.

Abstract 5: Astrology, Confirmation, Medium Quality

The purpose of this study is to evaluate the relative contributions of astrological determinants and astrological facets to the prediction of behavior. The authors compared five major astrological factors (sun sign, Ascendent, Medium Coeli, planets and major aspects) with five unfamiliar facets of astrology (minor aspects, lunar nodes, Asteroids, Uranian planets, Arabic parts) that constitute the astrological horoscope on their ability to predict 40 behavior criteria (e.g., alcohol consumption, willingness to share money). Both the major factors and the unfamiliar facets predicted substantial numbers of criteria, but the latter did noticeably better in that regard, even when the number of unfamiliar facet predictors was limited to the

number of the major predictors. Moreover, the criterion variance accounted for by the unfamiliar astrological facets often included large portions not predicted by the major astrological factors. The unfamiliar facets, therefore, were able to substantially increase the maximum prediction achieved by the major astrological factors. The results of this study are interpreted as supporting a more detailed approach to astrological assessment, one that goes beyond the interpretation of the major factors alone.

Method: A total of 47 undergraduate university students completed a questionnaire indicating their birth date, time and location of birth and a behavior report form. As mentioned, this study compares the major astrological factors with five more unfamiliar facets of astrology in terms of their criterion predictability. An astrologer was asked to rate the expected association between the diverse unfamiliar astrological facets and the criteria, and then chose the highest rated facets as our predictors of each criterion. For each subject a horoscope was calculated and an astrologer rated them separately for any astrological factor and facet in the 40 behavior criteria.

Results: In our first regression analyses all five major factor scales to predict each of the 40 behavior criteria were used. These five major predictors were able to account for significant amounts of variance in 15 of the 40 criterion variables. These behavior criteria were for example: Parties attended, $R=.41$, $p<.001$; Dieting behavior, $R=.38$, $p<.001$; Fraternity interest, $R=.31$, $p<.01$; Participation in sports, $R=.27$, $p<.05$; Alcohol consumption, $R=.27$, $p<.05$; Willingness to share money, $R=.25$, $p<.05$; Playing musical instruments, $R=.23$, $p<.05$. (R = multiple correlation of the predictors with criterion.)

In the second regression analyses the five unfamiliar facets of the astrological horoscope were used, rated by the astrologer, to predict each of the 40 behavior criteria. The five facets were able to account for significant amounts of variance in 17 of the 40 criterion variables. These behavior criteria were for example: Willingness to share money, $R=.48$, $p<.001$; Tobacco consumption, $R=.46$, $p<.001$; Religiosity rating, $R=.42$, $p<.001$; Alcohol consumption, $R=.34$, $p<.001$; Parties attended, $R=.31$, $p<.01$; Dating variety, $R=.28$, $p<.05$; Fraternity interest, $R=.26$, $p<.05$. (R = multiple correlation of the predictors with criterion.)

The five major astrological factors account for a substantial amount of variation in human behavior. Astrological interpretations of these factors, therefore, can be useful tools for predicting and understanding behavior. But there is also reliable astrological based variation in behavior that is not accounted for by the major factors. That additional variation is evident in the specific variance residing. Additional research in this area appears to be fruitful.

Abstract 6: Astrology, Confirmation, Low Quality

The purpose of this is to evaluate the relative contributions of astrological determinants to the prediction of behavior. The authors examined five coefficients of astrological factors (sun sign, Ascendent, Medium Coeli, planets and major aspects) on their ability to predict 40 behavior criteria (e.g., alcohol consumption, willingness to share money). The astrological personality factors predicted substantial numbers of criteria.

Method: A total of 27 undergraduate university students completed a questionnaire indicating their birth date, time and location of birth and a behavior report form. For each subject a Horoscope was conducted and an astrologer rated them separately for any astrological factor in the 40 behavior criteria.

Results: Correlations were computed with each coefficient of the astrological factors to predict each of the 40 behavior criteria. These astrological predictors were able to account for significant amounts of variance in 15 of the 40 criterion variables. These behavior criteria were for example: Parties attended with sun sign, $r=.41$, $p<.01$; Dieting behavior with major aspects, $r=.38$, $p<.01$; Fraternity interest with Ascendent, $r=.31$, $p<.01$; Participation in sports with planets, $r=.27$, $p<.05$; Alcohol consumption with sun sign, $r=.27$, $p<.05$; Willingness to share money with major aspects, $r=.25$, $p<.05$; Playing musical instruments with sun sign, $r=.23$, $p<.05$.

The five major astrological factors account for a substantial amount of variation in human behavior. Astrological interpretations of these factors, therefore, can be useful tools for predicting and understanding behavior. Additional research in this area appears to be fruitful.

Abstract 7: Big Five, Disconfirmation, High Quality

The purpose of this study is to evaluate the relative contributions of personality factors and personality facets to the prediction of behavior. The authors compared the Big Five factors of personality (Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience) with the facets or traits of personality that constitute those factors on their ability to predict 40 behavior criteria (e.g., alcohol consumption, willingness to share money). Both the broad factors and the narrow facets (the number of the narrow facet predictors was limited to the number of the broad factors) predicted no substantial numbers of criteria.

Method: A total of 187 subjects (97 female and 90 male university students with a mean age of 20.8) completed a set of self-report measures of personality (NEO-PI-R) and a behavioral report form. As mentioned, this study compares the Big Five Personality factors with five facets of those factors in terms of their criterion predictability. Judges were asked to rate the expected association between the diverse personality traits or facets and the criteria, and then the authors chose the five highest rated facets as predictors of each criterion.

Results (excerpts): In the first regression analyses all five NEO-PI-R factor scales to predict each of the 40 behavior criteria were used. These Big Five factor predictors were able to account for significant amounts of variance in only 2 of the 40 criterion variables. These behavior criteria were: Parties attended, $R1=.29$; $R0=.00$, $\%=6.4$, $p<.05$; and dieting behavior, $R1=.24$; $R0=.11$, $\%=6.2$, $p<.05$. ($R1$ = multiple correlation of sex and the predictors with criterion; $R0$ = multiple correlation of sex alone with the criterion; $\%$ = percentage of increase in criterion variance accounted for by $R1$ over $R0$.)

In the second regression analyses we the five facets of the NEO-PI-R factors to predict each of the 40 behavior criteria were used. The judges' five NEO-PI-R facet scales were able to account for significant amounts of variance in just 1 of the 40 criterion variables. This behavior criterion was: Fraternity interest, $R1=.22$, $R0=.04$, $\%=5.8$, $p<.05$. ($R1$ =

multiple correlation of sex and the predictors with criterion; $R0$ = multiple correlation of sex alone with the criterion; $\%$ = percentage of increase in criterion variance accounted for by $R1$ over $R0$.)

The Big Five factors of Personality do not account for a substantial amount of variation in human behavior. Measures of these factors may be useful tools for predicting and understanding behavior, but previous findings in this area are not supported in this study. There is also no reliable personality based variation in behavior that is not accounted for by the Big Five. Additional research in this area does not appear to be fruitful.

Abstract 8: Big Five, Disconfirmation, Medium Quality

The purpose of this study is to evaluate the relative contributions of personality factors and personality facets to the prediction of behavior. The authors compared the Big Five factors of personality (Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience) with the facets or traits of personality that constitute those factors on their ability to predict 40 behavior criteria (e.g., alcohol consumption, willingness to share money). Both the broad factors and the narrow facets (the number of the narrow facet predictors was limited to the number of the broad factors) predicted no substantial numbers of criteria.

Method: A total of 47 undergraduate university students completed a set of self-report measures of personality (NEO-PI-R) and a behavior report form.

As mentioned, this study compares the Big Five Personality factors with five facets of those factors in terms of their criterion predictability. The authors rated the expected association between the diverse personality traits or facets and the criteria, and then chose the highest rated facets as the predictors of each criterion.

Results: In the first regression analyses all five NEO-PI-R factor scales to predict each of the 40 behavior criteria were used. These Big Five factor predictors were able to account for significant amounts of variance in only 2 of the 40 criterion variables. These behavior criteria were: Parties attended, $R=.29$; $p<.05$; and dieting behavior, $R=.24$; $p<.05$. (R = multiple correlation of the predictors with criterion.)

In the second regression analyses the five facets of the NEO-PI-R factors to predict each of the 40 behavior criteria were used. The five NEO-PI-R facet scales were able to account for significant amounts of variance in just 1 of the 40 criterion variables. This behavior criterion was: Fraternity interest, $R=.22$, $p<.05$. (R = multiple correlation of the predictors with criterion.)

The Big Five factors of Personality do not account for a substantial amount of variation in human behavior. Measures of these factors may be useful tools for predicting and understanding behavior, but previous findings in this area are not supported in this study. There is also no reliable personality based variation in behavior that is not accounted for by the Big Five. Additional research in this area does not appear to be fruitful.

Abstract 9: Big Five, Disconfirmation, Low Quality

The purpose of this study is to evaluate the relative contributions of personality factors to the prediction of behavior. The authors examined the Big Five Factors of

personality (Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience) on their ability to predict 40 behavior criteria (e.g., alcohol consumption, willingness to share money). The Big Five personality factors predicted no substantial numbers of criteria.

Method: A total of 27 undergraduate university students completed a set of self-report measures of personality (NEO-PI-R) and a behavior report form.

Results: Correlations were computed with any of the five NEO-PI-R factor scales to predict each of the 40 behavior criteria. These Big Five factor predictors were able to account for significant amounts of variance in only 2 of the 40 criterion variables. These behavior criteria were: Parties attended with extraversion, $r=.29$; $p<.05$; and Dieting behavior with neuroticism, $r=.24$; $p<.05$.

The Big Five factors of Personality do not account for a substantial amount of variation in human behavior. Measures of these factors may be useful tools for predicting and understanding behavior, but previous findings in this area are not supported in this study. Additional research in this area does not appear to be fruitful.

Abstract 10: Big Five, Confirmation, High Quality

The purpose of this study is to evaluate the relative contributions of personality factors and personality facets to the prediction of behavior. The authors compared the Big Five factors of personality (Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience) with the facets or traits of personality that constitute those factors on their ability to predict 40 behavior criteria (e.g., alcohol consumption, willingness to share money). Both the broad factors and the narrow facets predicted substantial numbers of criteria, but the latter did noticeably better in that regard, even when the number of facet predictors was limited to the number of factor predictors. Moreover, the criterion variance accounted for by the personality facets often included large portions not predicted by the personality factors. The narrow facets, therefore, were able to substantially increase the maximum prediction achieved by the broad factors. The results of this study are interpreted as supporting a more detailed approach to personality assessment, one that goes beyond the measurement of the Big Five factors alone.

Method: A total of 187 subjects (97 female and 90 male university students with a mean age of 20.8) completed a set of self-report measures of personality (NEO-PI-R) and a behavioral report form. As mentioned, this study compares the Big Five Personality factors with five facets of those factors in terms of their criterion predictability. Judges were asked to rate the expected association between the diverse personality traits or facets and the criteria, and then the authors chose the five highest rated facets as predictors of each criterion.

Results (excerpts): In the first regression analyses all five NEO-PI-R factor scales to predict each of the 40 behavior criteria were used. These Big Five factor predictors were able to account for significant amounts of variance in 15 of the 40 criterion variables. These behavior criteria were for example: Parties attended, $R1=.41$, $R0=.00$, $\%=18.4$, $p<.001$; Dieting behavior, $R1=.38$, $R0=.13$, $\%=14.1$, $p<.001$; Fraternity interest, $R1=.31$, $R0=.11$, $\%=10.5$, $p<.01$; Participation in sports, $R1=.27$, $R0=.09$, $\%=8.3$, $p<.05$; Alcohol consumption, $R1=.27$, $R0=.06$, $\%=7.9$, $p<.05$; Willingness to share money, $R1=.25$, $R0=$

$.11$, $\%=7.4$, $p<.05$; Playing musical instruments, $R1=.23$, $R0=.00$, $p<.05$, $\%=7.1$. ($R1$ = multiple correlation of sex and the predictors with criterion; $R0$ = multiple correlation of sex alone with the criterion; $\%$ = percentage of increase in criterion variance accounted for by $R1$ over $R0$.)

In the second regression analyses the five facets of the NEO-PI-R factors to predict each of the 40 behavior criteria were used. The judges' five NEO-PI-R facet scales were able to account for significant amounts of variance in 17 of the 40 criterion variables. These behavior criteria were for example: Willingness to share money, $R1=.48$, $R0=.16$, $\%=20.1$, $p<.001$; Tobacco consumption, $R1=.46$, $R0=.00$, $\%=19.8$, $p<.001$; Religiosity rating, $R1=.42$, $R0=.11$, $\%=17.6$, $p<.001$; Alcohol consumption, $R1=.34$, $R0=.03$, $\%=15.3$, $p<.001$; Parties attended, $R1=.31$, $R0=.00$, $\%=12.4$, $p<.01$; Dating variety, $R1=.28$, $R0=.09$, $\%=8.9$, $p<.05$; Fraternity interest, $R1=.26$, $R0=.04$, $\%=8.4$, $p<.05$. ($R1$ = multiple correlation of sex and the predictors with criterion; $R0$ = multiple correlation of sex alone with the criterion; $\%$ = percentage of increase in criterion variance accounted for by $R1$ over $R0$.)

The Big Five factors of Personality account for a substantial amount of variation in human behavior. Measures of these factors, therefore, can be useful tools for predicting and understanding behavior. But there is also reliable personality based variation in behavior that is not accounted for by the Big Five. This additional variation is evident in the specific variance residing. Additional research in this area appears to be fruitful.

Abstract 11: Big Five, Confirmation, Medium Quality

The purpose of this study is to evaluate the relative contributions of personality factors and personality facets to the prediction of behavior. The authors compared the Big Five factors of personality (Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience) with the facets or traits of personality that constitute those factors on their ability to predict 40 behavior criteria (e.g., alcohol consumption, willingness to share money). Both the broad factors and the narrow facets predicted substantial numbers of criteria, but the latter did noticeably better in that regard, even when the number of facet predictors was limited to the number of factor predictors. Moreover, the criterion variance accounted for by the personality facets often included large portions not predicted by the personality factors. The narrow facets, therefore, were able to substantially increase the maximum prediction achieved by the broad factors. The results of this study are interpreted as supporting a more detailed approach to personality assessment, one that goes beyond the measurement of the Big Five factors alone.

Method: A total of 47 undergraduate university students completed a set of self-report measures of personality (NEO-PI-R) and a behavioral report form.

As mentioned, this study compares the Big Five Personality factors with five facets of those factors in terms of their criterion predictability. The authors rated the expected association between the diverse personality traits or facets and the criteria, and then chose the five highest rated facets as the predictors of each criterion.

Results: In the first regression analyses all five NEO-PI-R factor scales to predict each of the 40 behavior criteria were used. These Big Five factor predictors were able to account for significant amounts of variance in 15 of the 40 criterion variables.

These behavior criteria were for example: Parties attended, $R=.41$, $p<.001$; Dieting behavior, $R=.38$, $p<.001$; Fraternity interest, $R=.31$, $p<.01$; Participation in sports, $R=.27$, $p<.05$; Alcohol consumption, $R=.27$, $p<.05$; Willingness to share money, $R=.25$, $p<.05$; Playing musical instruments, $R=.23$, $p<.05$. (R = multiple correlation of the predictors with criterion.)

In the second regression analyses the five facets of the NEO-PI-R factors to predict each of the 40 behavior criteria were used. The five NEO-PI-R facet scales were able to account for significant amounts of variance in 17 of the 40 criterion variables. These behavior criteria were for example: Willingness to share money, $R=.48$, $p<.001$; Tobacco consumption, $R=.46$, $p<.001$; Religiosity rating, $R=.42$, $p<.001$; Alcohol consumption, $R=.34$, $p<.001$; Parties attended, $R=.31$, $p<.01$; Dating variety, $R=.28$, $p<.05$; Fraternity interest, $R=.26$, $p<.05$. (R = multiple correlation of the predictors with criterion.)

The Big Five factors of Personality account for a substantial amount of variation in human behavior. Measures of these factors, therefore, can be useful tools for predicting and understanding behavior. But there is also reliable personality based variation in behavior that is not accounted for by the Big Five. This additional variation is evident in the specific variance residing. Additional research in this area appears to be fruitful.

Abstract 12: Big Five, Confirmation, Low Quality

The purpose of this study is to evaluate the relative contributions of personality factors to the prediction of behavior. The authors examined the Big Five Factors of personality (Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience) on their ability to predict 40 behavior criteria (e.g., alcohol consumption, willingness to share money). The Big Five personality factors predicted substantial numbers of criteria.

Method: A total of 27 undergraduate university students completed a set of self-report measures of personality (NEO-PI-R) and a behavioral report form.

Results: Correlations were computed with any of the five NEO-PI-R factor scales to predict each of the 40 behavior criteria. These Big Five factor predictors were able to account for significant amounts of variance in 15 of the 40 criterion variables. These behavior criteria were for example: Parties attended with extraversion, $r=.41$, $p<.01$; Dieting behavior with neuroticism, $r=.38$, $p<.01$; Fraternity interest with extraversion, $r=.31$, $p<.01$; Participation in sports with openness to experience, $r=.27$, $p<.05$; Alcohol consumption with extraversion, $r=.27$, $p<.05$; Willingness to share money with agreeableness, $r=.25$, $p<.05$; Playing musical instruments with conscientiousness, $r=.23$, $p<.05$.

The Big Five factors of Personality account for a substantial amount of variation in human behavior. Measures of these factors, therefore, can be useful tools for predicting and understanding behavior. Additional research in this area appears to be fruitful.

References

- Armstrong, J. S. (1996). Publication of research on controversial topics: the early acceptance procedure. *International Journal of Forecasting*, 12, 299–302.

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum.
- Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representations of human judgment* (pp. 17–52). New York: Wiley.
- Epstein, W. M. (2004). Confirmation bias and the quality of the editorial processes among American social work journals. *Research on Social Work Practice*, 14, 450–458.
- Goodstein, L. D., & Brazis, K. L. (1970). Psychology of scientist: XXX. Credibility of psychologists: an empirical study. *Psychological Reports*, 27, 835–838.
- Greenwald, A. G., Pratkanis, A. R., Leippe, M. R., & Baumgardner, M. H. (1986). Under what conditions does theory obstruct research progress? *Psychological Review*, 93, 216–229.
- Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological Bulletin*, 135, 555–588.
- Horrobin, D. F. (1990). The philosophical basis of peer review and the suppression of innovation. *Journal of the American Medical Association*, 263, 1438–1441.
- Kochler, J. J. (1993). The influence of prior beliefs on scientific judgments of evidence quality. *Organizational Behavior and Human Decision Processes*, 56, 28–55.
- Krems, J. F., & Zierer, C. (1994). Sind Experten gegen kognitive Täuschungen gefeit? Zur Abhängigkeit des confirmation bias von Fachwissen. [Are experts immune to cognitive bias? The dependence of confirmation bias on specialist knowledge]. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 41, 98–115.
- 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, 2098–2109.
- Maccoun, R. J. (1998). Biases in the interpretation and use of research results. *Annual Review of Psychology*, 49, 259–87.
- Mahoney, M. J. (1977). Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research*, 1, 161–175.
- Moss, S., & Butler, D. C. (1978). The scientific credibility of ESP. *Perceptual and Motor Skills*, 46, 1063–1079.
- Nickerson, R. S. (1998). Confirmation bias: a ubiquitous phenomenon in many guises. *Review of General Psychology*, 2, 175–220.
- Oswald, M. E., & Grosjean, S. (2004). Confirmation bias. In R. Pohl (Ed.), *Cognitive illusions. A handbook on fallacies and biases in thinking, judgement and memory* (pp. 79–96). Hove: Psychology Press.
- Paunonen, S. V., & Ashton, M. C. (2001). Big Five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology*, 81, 524–539.
- Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, 32, 880–892.
- Van Ophuysen, S. (2006). Comparison of diagnostic decisions between novices and experts: the example of school career recommendation. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie*, 38, 154–161.