Podlesny, J.R., & Raskin, D.C. (1977). Physiological measures and the detection of deception. *Psychological Bulletin, 84*, 782-799.

Reeder, G.D., & Brewer, M.B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review, 86*, 61-79.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 10, pp. 173-220). New York: Academic Press

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.

Schachter, S. (1964). The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1, pp. 49-80). New York: Academic Press.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional states. *Psychological Review, 69*, 379-399.

Semin, G.R., & Fiedler, K. (1991). The linguistic category model, its bases, applications, and range. In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology,* (Vol. 2, pp. 1-30). Chichester: Wiley.

Semin, G.R., & Marsman, G. (1994). On the 'multiple inference-inviting properties' of interpersonal verbs: Event instigation, dispositional inferences, and implicit causality. *Journal of Personality and Social Psychology, 67*, 836 - 849.

Skowronsky, N.J., & Carlston, D.E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin, 105*, 131-142.

Szucko, J.J., & Kleinmuntz, B. (1981). Statistical versus clinical lie detection. *American Psychologist, 36*, 488-496.

Taylor, S.E., & Fiske, S.T. (1978). Salience, attention, and attribution: Top of the head phenomena. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 249-288). San Diego, CA: Academic Press.

Turner, R.E., Edgley, C., & Olmstead, G. (1975). Information control in conversation: Honesty is not always the best policy. *Kansas Journal of Sociology, 11*, 69-89.

Tversky, A., & Kahnemann, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.

Uleman, J.S. (1987). Consciousness and control: The case of spontaneous trait inferences. *Personality and Social Psychology Bulletin, 13*, 337-354.

Watson, D. (1982). The actor and the observer: How are their perceptions of causality divergent? *Psychological Bulletin, 92*, 682-700.

Zuckerman, M., DePaulo, B.M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology, 14*, 1-59.

Zuckerman, M., Fischer, S.A., Osmun, R.W. Winkler, B.A., & Wolfson, L.R. (1987). Anchoring in lie detection revisited. *Journal of Nonverbal Behavior, 11*, 4-12.

Zuckerman, M., Koestner, R., Colella, M.J., & Alton, A.O. (1984). Anchoring in the detection of deception and leakage. *Journal of Personality and Social Psychology, 47*, 301-311.

Research underlying this chapter was supported by a grant from the Deutsche Forschungsgemeinschaft.

Suggestibility and Anchoring

Rüdiger F. Pohl

■ *In their attempt to estimate the numerical answers to difficult questions, persons can be easily influenced by presenting them estimates of other persons. One explanation of this so-called anchoring effect assumes that the suggested anchor value selectively activates one's knowledge base, thus biasing any future estimation process. This process may, however, depend on the individual's suggestibility. This paper presents an experiment that explores the influence of suggestibility on anchoring. An anchor was found to be more influential when (a) it came from an expert rather than from a layman and when (b) it was a precise value rather than an interval. Additionally, (c) an anchor was maximally influential when it was subjectively rated as "good", while it had no effect at all when it was rated as "poor". However, (d) the individual suggestibility was not related to the amount of anchoring.*

Anchoring effect and hindsight bias are two phenomena that have been amply studied and documented (see, e.g., Jacowitz & Kahneman, 1995; Pohl, 1998). *Anchoring* seems to occur whenever someone is asked to estimate an unknown quantity in the light of some (sometimes even arbitrary) numerical information, that is, an "anchor". Following a classical study (Tversky & Kahneman, 1974), participants usually first judge whether the unknown quantity is larger or smaller than the given anchor before generating an exact estimate. In the *hindsight* paradigm, participants receive the solution as answer to a difficult question and are then asked to generate an estimate "as if they had not known the solution" (see, e.g., Davies, 1992).

Both experimental techniques lead to similar observations in that participants' estimates are similarly biased towards the anchor or the solution, respectively. Another finding is that both phenomena are rather robust and are thus difficult to overcome (Pohl & Hell, 1996; Wilson, Houston, Etling, & Brekke, 1996). And finally, similar cognitive processes are assumed to be responsible for both biases (Pohl & Eisenhauer, 1997;

Strack & Mussweiler, 1997). Because of this more but superficial similarity, the term "anchoring" will be used in the following to encompass both experimental paradigms (Pohl, 1995).

In order to explain anchoring effects, one must distinguish between two sources. These are (1) characteristics of the experimental situation and (2) individual differences. The following discussion of these factors will concentrate on such experimental and individual characteristics that are related to suggestibility. A review of other influencing factors can be found elsewhere (Hawkins & Hastie, 1990).

Anchoring may be understood as a special case of semantic *priming* (Jacowitz & Kahneman, 1995; Strack & Mussweiler, 1997). However, while semantic priming is considered to proceed automatically, anchoring appears to include an evaluative stage, in which certain characteristics of the anchor value may become effective. These are (among others) the anchor's source and its numerical plausibility.

From the suggestibility literature it is known that a more authoritative *source* may have a stronger distorting effect than a less authoritative one (e.g., Toglia, Ross, Ceci, & Hembrooke, 1992). Similarly, models of persuasion and attitude change like the "Heuristic-Systematic Model" (Chaiken, Liberman, & Eagly, 1989) acknowledge that the source's credibility may act as a heuristic cue thus influencing someone's attitude (Chaiken & Maheswaran, 1994). In terms of the "Elaboration Likelihood Model" (Petty & Cacioppo, 1986), the source of an information may serve as a peripheral cue leading to a biased elaboration of that information.

Only a few studies employed an anchor-source manipulation. Switzer and Sniezek (1991) varied the relevance of the apparent source of their anchor values, but found no reliable differences in the resulting anchor effects. Similarly, Plous (1989, Exp. 6) labeled the anchor value as originating either from other people or from an expert, but found no moderating influence on anchoring. In an unpublished study (cited in Jacowitz & Kahneman, 1995, p. 1164), the manipulation of the credibility of an anchor did not change the effect. Using a hindsight design, Pohl (1998) labeled the anchor values either as being from another participant from a previous study or as the solution to the question. Again, anchoring was identical under both conditions. All these findings would argue against the existence of an evaluation stage. When, however, the subjective plausibility of each anchor was taken into account (Pohl, 1998), subjectively implausible anchors assumed to come from another participant lost their distorting influence, while the solution-labeled anchors still led to hindsight bias. This result supports the assumptions of an evaluation stage and one might hesitate to interpret the other findings as evidence against such a stage. Maybe, the effect was simply overlooked, because the data were summed across subjectively plausible and implausible anchor values.

The anchor's numerical *plausibility* should thus be considered as a second influencing factor. Most likely, each anchor value is compared to the participant's item-specific knowledge base (Pohl & Eisenhauer, 1997; Strack & Mussweiler, 1997). One characteristic of the individual knowledge base is the amount of knowledge that enables the participant to generate a more or less restricted range of plausible values. Only if the anchor value falls within that range, an anchor will be considered plausible. If anchoring is more pronounced in this case, the explanation of anchoring must include some evaluative stage.

Most studies used a predetermined degree of plausibility (e.g., by evaluating the distribution of unbiased estimates). Again, in the domain of suggestibility, a more plausible suggestion showed a larger influence than a less plausible one (e.g., Loftus, 1979; Read & Bruce, 1984). Using an anchoring design, Chapman and Johnson (1994) reported a reduced anchoring effect for implausible anchors, while Strack and Mussweiler (1997, Exp. 3) found no significant difference between plausible and implausible anchors. Similarly, Pohl (1998) found no effect of the anchor's "objective" plausibility. But when the data were split according to the anchor's subjective plausibility, the above mentioned result emerged: subjectively implausible estimates from another participant lost their distorting influence, while the solution-labeled anchors still led to a hindsight bias - even if they were considered implausible.

In addition to (or in interaction with) the anchor's source and plausibility, the individual *suggestibility* may play a role as well. Unfortunately, however, up to now all studies in the field of cognitive illusions have considered suggestibility as a dependent variable rather than an independent one. That is, the larger the individual anchoring effect, hindsight bias, or misinformation effect, the larger the individual suggestibility (see, e.g., Rhodes & Wood, 1992). The following experiment took the opposite approach. The question was whether individual suggestibility (as measured in an independent test) might predict the individual amount of anchoring.

Aside from the influencing factors described above, it remains to be explained how anchoring comes about in the first place. Wilson et al. (1996) noted that "anchoring processes themselves are poorly understood" (p. 387). A recently proposed *cognitive process model* named "SARA" (Pohl & Eisenhauer, 1997) was set out to remedy the situation (see also Pohl, 1998). The acronym SARA stands for *"Selective Activation, Reconstruction, and Anchoring"*. The model encompasses a detailed description of all the processes assumed responsible for producing anchoring and hindsight bias. In addition, the model has been implemented as a computer model that successfully simulated empirically observed data.

SARA is a semantic memory model including assumptions about someone's knowledge base, about encoding and retrieval processes, and about forgetting. SARA's general architecture is based on "SAM" - the *Search of Associative Memory* model (Raaijmakers & Shiffrin, 1980; Shiffrin & Raaijmakers, 1992). According to SARA, anchored reconstructions result from a *selective activation* of one's item-specific knowledge base (Hawkins & Hastie, 1990; Leary, 1982; Pohl & Gawlik, 1995; Slovic & Fischhoff, 1977; Strack & Mussweiler, 1997). This activation is governed by the anchor value and is considered as being selective, because information that is more similar to (or consistent with) the anchor will receive more activation than other information (Kahneman & Miller, 1986). After selective activation, the probability of retrieving a certain piece of information from one's knowledge base has changed. As a consequence, any attempt to generate or reconstruct an "unbiased" estimate is bound to fail. Most probably, the resulting estimate will be biased towards the anchor value.

SARA might be able to simulate the impact of the anchor's source and its plausibility by assigning different weights to the retrieval cues (e.g., the anchor values) that are used to search one's memory. Hopefully then, SARA will provide us with a more satisfactory explanation of recall and judgment distortions (like hindsight bias and anchoring) than the mainly descriptive models that have been proposed to date (Jacowitz & Kahneman, 1995). First, however, it remains to be tested what influence the source and the plausibility of an anchor actually have.

Method

Design

The experiment was based on a three-factor design. (a) The *source* of the anchor value was labeled either as being a layman ("a randomly selected person passing by on the street") or as being an expert ("professor of psychology"). (b) The *precision* of the anchor value was either high (exact value, e. g., "45%") or low (interval, e. g., "39-51%"). (c) The *position* of the anchor value was either above or below the mean unbiased estimates that had been collected in a previous experiment. The first variable constituted a grouping factor, while the other two were introduced as repeated-measures factors.

Materials

The materials consisted of a questionnaire and two personality scales. The *questionnaire* included 50 items asking about difficult numerical quantities from all areas of psychology (e. g., "What percentage of a list of 13 syllables is retained after 8 hours?" or "What percentage of adult Americans can be hypnotized without practice?"). All questions had to be answered with a value between 0% and 100%. Each question was accompanied by an anchor value that was labeled as being another person's estimate. In the introduction, these estimates were classified as "possibly good, but not necessarily good" estimates that had been collected in a previous experiment. In one group, these estimates were said to come from a *layman* ("person passing by"), while in the other they were said to come from an *expert* ("professor of psychology"). The precision and position of the anchor values were systematically varied across participants so that each item was presented equally often in each condition.

The *anchor values* in the precise-anchor condition were computed by either adding or subtracting 21 percentage points from the mean unbiased estimates that had been collected in a previous experiment. If, for example, the mean unbiased estimate was 45%, the low anchor would be 34% and the high anchor 66%. In the interval condition, an interval of 12 percentage points was constructed around the precise anchor values, that is, 28-40% in the low-anchor and 60-72% in the high-anchor condition.

The two personality scales focussed on individual differences in suggestibility and dogmatism. The *suggestibility* test was based on a technique reported by Burger (1971). The test consisted of a complex photograph (of a car advertisement) and a 16-item questionnaire including 8 questions with false suggestions (e. g., "Was the dog in the foreground a dachshund or a cocker spaniel?" when in fact there was no dog in the picture). The *dogmatism* was assessed with a German translation of the short form of the Rokeach (1960) Dogmatism Scale. This scale included 18 items that had to be rated on a 6-point rating scale (from -3 = "I absolutely disagree" to +3 = "I absolutely agree" without a zero-point).

Participants

Eighty students (51 female, 29 male; between 19 and 44 years old with a mean age of 23.6 years) from different faculties of the Catholic University of Eichstätt (Germany) volunteered to participate in the experiment. They were either paid DM 12 or received course credit.

Procedure

The experiment was run in small groups from one to five participants. The 80 participants were randomly assigned to one of the two source-label groups (layman or expert).

Each participant first completed the 50-item questionnaire and then received the personality scales. Regarding the questionnaire, the tasks were as follows. After reading the question and the other person's estimate (i. e., the anchor value), the participant was asked to rate the quality of the given estimate on a dichotomous scale as "rather good" or "rather poor". Then the participant generated his or her own estimate together with a 7-point confidence rating (from 1 = "very sure" to 7 = "very unsure"). After completion of the questionnaire, the two personality tests were run.

Participants were first presented the photograph for 5 seconds and were then asked to answer the 16 questions. The number of erroneously accepted false alternatives was taken as an index of suggestibility. Finally, participants rated the 18 items of the dogmatism scale.

The whole experiment lasted between 45 and 60 minutes, depending on the participant's individual pace. There was no time limit.

Analysis

From a total of 4,000 observations (= 80 participants x 50 items), one estimate, five anchor ratings, and 23 confidence ratings were missing. As dependent measure, the *mean deviation D* was computed for each participant *i* by averaging all differences between a participant's estimate E_{ij} and the corresponding mean unbiased estimate $\overline{E}_j$ (collected in a previous experiment) across all m items in each experimental condition:

$$D_i = \sum_{j=1}^{m} (E_{ij} - \overline{E}_j)$$

A positive value of D would indicate that the participant's estimates were generally higher than the unbiased estimates, while a negative D would indicate that they were generally lower than the unbiased estimates. In addition, an *anchor index A* was computed using the difference between the mean deviations D in the two anchor-direction conditions (i. e., "-21" and +21") and by dividing this difference by the total distance between high and low anchor (i. e., 42). This index is identical to the slope of the regression line connecting the mean estimates for high and low anchors (cf. Jacowitz & Kahneman, 1995; Pohl, 1995):

$$Ai = (Di^{+21} - Di^{-21}) / 42 = (Ei^{+21} - Ei^{-21}) / 42$$

A positive index would indicate an anchoring effect, an index of zero the absence of any effect, and a negative index a contrast effect.

While the mean deviation D will serve to illustrate the results, the anchor index A was preferred for analyzing the data in several ANOVAs. The level of significance was set to $\alpha = .05$ for all analyses.

Results

Anchoring

Overall, the participants' estimates deviated 7.8 percentage points from the unbiased mean into the direction of the anchor value indicating a large anchoring effect (Fig. 1).

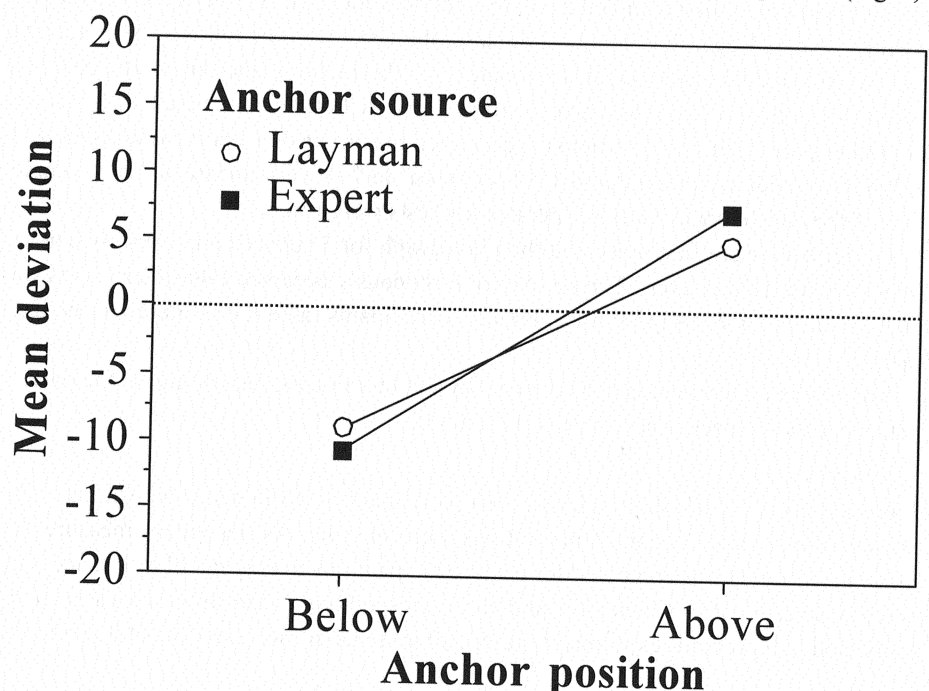

Figure 1: *Mean deviations of observed estimates from unbiased estimates (zero line).*

Table 1. *Mean anchor indices for all experimental conditions*

	Exact anchor	Interval anchor	Mean
Expert source	.44	.40	.42
Layman source	.34	.34	.34
Mean	.39	.37	.38

This deviation was similar for male (7.6) and female (7.9) participants, $t\,(78)=-.388$, $p=.70$. In terms of the anchor index A, the participants' estimates deviated 37% (= 15.5/42) of the distance between unbiased estimate and anchor value. The individual anchoring indices ranged from .06 to .82 with a mean of .38 and a standard deviation of .17.

Analyzing the anchor indices A, a 2x2 ANOVA with the factors source and precision led to the following results. The anchor effect was significantly larger for an expert source (i.e., .42) than for a layman source (i.e., .34), $F(1,78) = 5.8$, $MSE = .05$. The precision of the anchor played no significant role, neither as a main effect nor in interaction with the source, both $Fs < 1$. The mean anchor indices for all experimental conditions are given in Table 1.

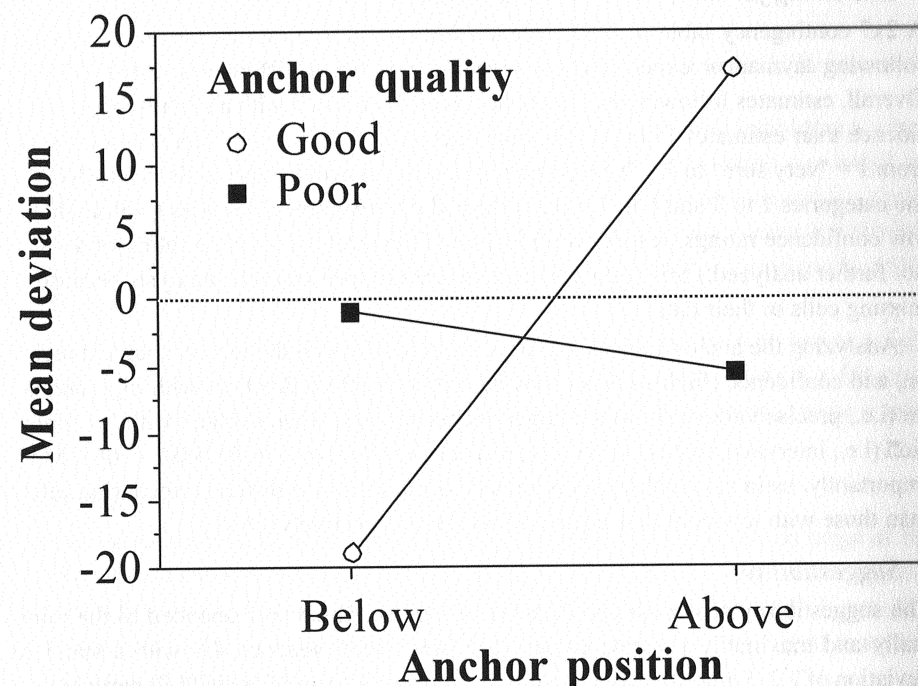

Figure 2: *Mean deviations of observed estimates from unbiased estimates (zero line) for anchors rated as "rather good" vs. anchors rated as "rather poor".*

Anchoring for good and poor anchors

Overall, exactly 50% of all anchors were rated as "rather good" and 49.9% as "rather poor" (5 ratings were missing). Analyzing the frequencies of anchors that were rated as "rather good" by means of a 2x2x2 ANOVA with the factors source, direction, and precision revealed all three main effects to be significant, but none of the interactions. Anchors from the expert source were more frequently rated as good than those from the layman source (52.5% vs. 47.5%, resp.), $F(1,78) = 4.183$, $MSE = 7.471$. In addition, interval anchors were more frequently rated as good than precise anchors (53.2% vs. 46.8%, resp.), $F(1,78) = 15.446$, $MSE = 1.250$. And finally, anchors above the unbiased mean were rated more frequently as good than those below the unbiased mean (52.4% vs. 47.6%, respectively), $F(1,78) = 8.082$, $MSE = 1.250$.

The anchor indices A for good and poor anchors were analyzed in a 2x2x2 ANOVA with the factors source, precision, and subjective anchor quality (good vs. poor). The anchoring effect was significantly larger for subjectively good than for subjectively bad anchors (i.e., .86 vs. -.11, respectively), $F(1,78) = 1325$, $MSE = .057$. And in regard of the anchor's precision, the anchoring effect was larger for precise values than for intervals (i.e., .41 vs. .34, respectively), $F(1,78) = 7.56$, $MSE = .051$. No other effect emerged.

Anchoring for more or less confident estimates

A 2x7 contingency table of the frequencies of confidence ratings for one's estimates following layman or expert anchors revealed significant differences, $\chi^2(6) = 23.17$. Overall, estimates following expert anchors were associated with a slightly higher confidence than estimates following layman anchors (i.e., 3.6 vs. 3.4 on a 7-point scale from 1 = "very sure" to 7 = "very unsure"). In order to avoid empty cells in an ANOVA, the categories 1 to 3 and 5 to 7 were collapsed into two new categories (with high and low confidence ratings, respectively). (The middle category - with a rating of 4 - was not further analyzed.) Still, 30 participants were dropped from the analysis because of missing cells in their data.

Analyzing the anchor indices A in a 2x2x2 ANOVA with the factors source, precision, and confidence (high vs. low) showed only one effect. Anchors with high precision (i.e., precise values) showed a larger anchoring effect than anchors with low precision (i.e., intervals), that is, .43 vs. .34, respectively, $F(1,48) = 6.80$, $MSE = .061$. More importantly, estimates with high confidence did not show a different anchoring effect than those with low confidence (i.e., .37 vs. .39, respectively), $F<1$.

Suggestibility

The suggestibility scores ranged from 0 to 8 points, which corresponded to the minimally and maximally possible values. The mean score reached 4.8 with a standard deviation of 2.2. Suggestibility was similar for male (4.6) and female (4.9) participants, $t(78) = -.658$, $p = .51$. Two methods were employed to assess the influence of suggestibility on the anchoring effect.

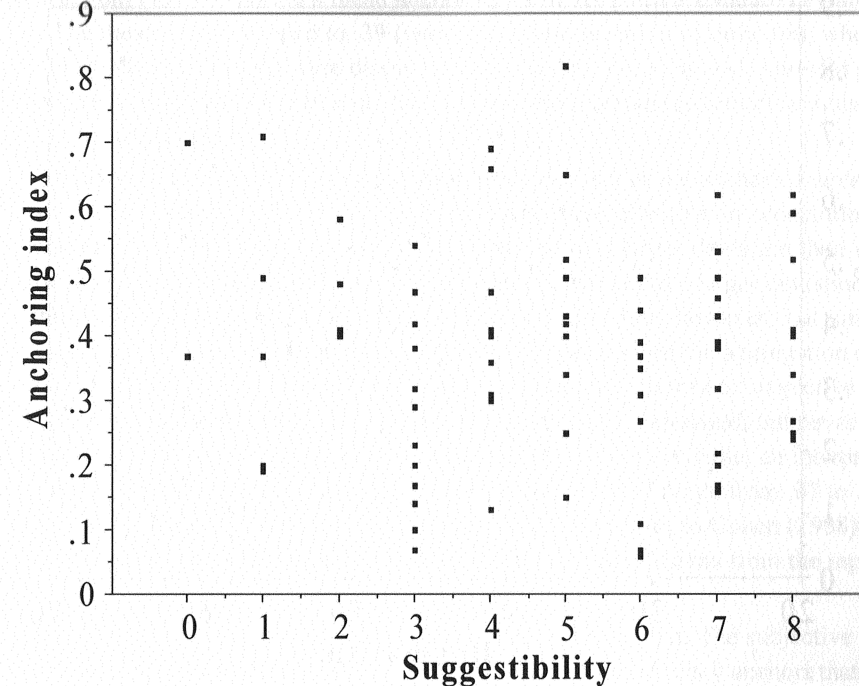

Figure 3: *Regression plot of individual anchoring index and suggestibility scores.*

First, the correlation between individual suggestibility scores and individual anchoring effects was computed (see Fig. 3). The product-moment correlation was found to be -.05 ($N = 80$), which was not significantly different from zero (one-tailed t-test). Computing separate correlation scores for layman and expert anchors (each $N = 40$) led to scores of -.20 and .02, respectively. Again, these scores were not significantly different from zero (one-tailed t-tests).

In a second attempt, the participants were divided into two groups with low and high suggestibility scores, that is, scores from 0 to 4 ($N = 35$) and from 5 to 8 ($N = 45$), respectively. The mean anchoring index was found to be .38 in both groups (making any statistical test obsolete).

Dogmatism

The individual dogmatism scores ranged from 26 to 69 (out of a maximum range of 18 to 108) with a mean score of 50.6 and a standard deviation of 9.2. Again, there was no difference between male (49.4) and female (51.2) participants, $t(78) = -.813$, $p = .42$. Corresponding to the analysis of suggestibility, two methods were employed to assess the influence of dogmatism.

First, the correlation between individual dogmatism scores and individual anchoring effects was computed (see Fig. 4). The product-moment correlation between dogma-

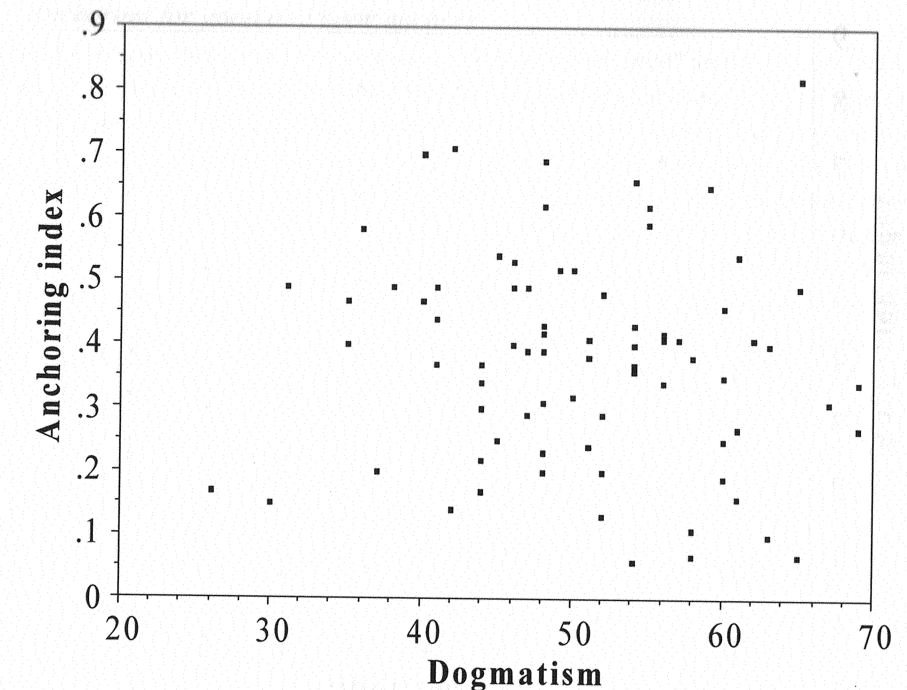

***Figure 4:** Regression plot of individual anchoring index and dogmatism scores.*

tism and anchoring was figured to be -.08 ($N = 80$), which was not significantly different from zero (one-tailed t-test). This result was apparently not caused by any extreme data point.

Splitting the correlation score for the layman and the expert-source group (with each $N = 40$), led to negative correlation scores in both groups, that is, -.22 and -.09, respectively. Both scores were not significantly different from zero (one-tailed t-tests).

As a second approach, only participants with extreme dogmatism scores taken from the highest ($N = 20$) and lowest ($N = 20$) quartile of the dogmatism distribution were considered. Individuals with high and low dogmatism showed mean anchoring indices of .33 and .38, respectively. These indices were not significantly different, $t(38) = .852$, $p = .40$.

Discussion

As in many previous studies, the estimates of participants were influenced by the presentation of numerical anchor values, thus replicating the *anchoring effect*. The mean estimates deviated (from unbiased means collected in an earlier experiment) towards the anchor value. The overall anchor index (Jacowitz & Kahneman, 1995; Pohl, 1995) was found to be .37, which means that the participants' estimates deviated 37% of the distance between unbiased estimates and anchor value. In his summary of 22 anchor studies, Pohl (1995) reported a mean anchor index of .41 (with a standard deviation of .16). The mean index dropped to .39 (with a standard deviation of only .08), when the two most extreme findings were discarded. The present result fits well into that picture and, thus, suggests not only a qualitative, but also a certain quantitative robustness of the anchoring effect.

Both experimental manipulations of the experiment, that is, the anchor's *source* (layman vs. expert) and its *precision* (high vs. low), were effective in moderating the anchoring effect. Generally, an expert's estimate led to a larger deviation than a layman's estimate and a precise anchor (i.e., an exact value) led to a larger deviation than an imprecise one (i.e., interval values). These effects were, however, not simultaneously present in any of the statistical analyses. Depending on the aggregation of the data (according to the two main factors or to the additional factors of subjective plausibility and of subjective confidence), one or the other effect emerged, but never both. This suggests that both effects were rather small and not very reliable, an interpretation that is supported by the effect sizes that were computed to range from .07 to .14 in these cases. Such effect sizes are considered "small" according to Cohen (1988). One should therefore proceed with some caution in drawing conclusions from the reported findings.

Another effect, however, was unexpectedly large and clearcut. The subjective *plausibility* determined whether or not an anchoring effect occurred. Only anchors that were subjectively considered plausible led to an anchoring effect (with an index of .86), while anchors that were considered implausible led to no effect at all (with an index of -.11). This finding strongly supports the notion that an evaluative stage must be part of the anchoring process (Pohl, 1998). Only if the anchor passes this preliminary test, it will exert its biasing influence. The criteria for acceptance are presumably the anchor's numerical value, its source, its precision and maybe other characteristics. In the present experiment, the numerical plausibility of the anchor clearly had the largest consequences for anchoring, while the source and the precision of the anchor played only minor roles. Maybe that difference in effectiveness is also the reason why there wasn't any interaction of the plausibility effect with the anchor's source or its precision (a finding that is different from what Pohl, 1998, found). Another interpretation of the missing interaction is that the anchor's source and its precision were already taken into account in the evaluation stage. One result supporting this idea was that the number of subjectively "good" anchors was higher for the expert source than for the layman source. The same argument, however, can not be applied to the precision factor. Here, the number of subjectively "good" anchors was higher for imprecise anchors (i.e., intervals) than for precise ones (i.e., exact values), but the latter showed a larger anchoring effect.

Another consequence of the strong plausibility effect found here is that any anchoring effect may actually be a compound of items with a strong ("true") effect and items without one. If plausible and implausible anchors are about equally frequent in a study and if, moreover, their effects are similar to those found here (i.e., about 80% and 0%, respectively), then the observed mean anchoring of around 40% (Pohl, 1995) simply represents an artifact.

The *confidence* in one's own estimates had apparently no diagnostic value in differentiating between more or less biased estimates. This result is in strong contrast with those reported by Jacowitz and Kahneman (1995). In their own study they found a significant negative correlation of -.68 between confidence and anchoring. They additionally cite two unpublished sources (p. 1164) that also support the notion that the size of the anchoring effect varies inversely with the participant's confidence in their estimates. One explanation of these divergent findings might be that, in the present experiment, the original 7-point confidence ratings were collapsed into two categories of high and low confidence (with the middle category left out). This procedure, of course, led to a reduction of sensitivity. Moreover, only the mean anchoring scores for high and low confidence were subjected to an ANOVA, thus possibly further obscuring any effect that might have been present in the raw data. Computing the correlation coefficient from the raw data, however, confirmed the reported ANOVA results. The overall correlation between confidence and deviation into the anchor's direction was found to be r = -002 (N = 3,977), with r = -011 for the layman-source group (N = 1,985) and r =.012 for the expert-source group (N = 1,992). All three scores were not significantly different from zero (one-tailed t-tests). Thus it remains unclear why Jacowitz and Kahneman (1995) found such a strong relationship.

Less controversial than the confidence effects were the results on individual *suggestibility* and *dogmatism*. Both measures showed no relation whatsoever to the amount of anchoring (see Figures 3 and 4). One might, of course, argue that a sample of participants with more extreme degrees of suggestibility or dogmatism might have shown some effect. However, the present participants' scores already covered most of the appropriate scales leaving little room for more extreme scores. Another problem, though, could be more important here. Most studies that set out to find a correlation between some sort of biased judgment or recall and certain personality characteristics failed to do so or uncovered only small relationships (see, e.g., Campbell & Tesser, 1983). One reason for this might be that, on an individual basis, biasing (like anchoring, hindsight or misinformation) is not a reliable phenomenon. Pohl (2000) recently computed split-half reliability scores on the data of 29 such studies and found a mean reliability of only r = 072, which is not significantly different from zero, $t(27)$ = .375, p =.355 (one-tailed). Out of the 29 studies, only two showed a substantial and statistically significant reliability (of .28 and .50), while each of the 27 other scores was not significantly different from zero. But if these phenomena are unreliable when regarding the individual, then, of course, any search for possibly responsible personality characteristics becomes immediately obsolete.

In terms of the computer model SARA (Pohl & Eisenhauer, 1997; Pohl, 1998), the observed effects could be explained and simulated as follows. The anchor's plausibility and additionally its source and its precision--if they were found to reliably influence anchoring--would all change the weights associated with the anchor as a retrieval cue. Thus, any piece of knowledge from one's knowledge base associated with that anchor would receive a higher or smaller probability to be retrieved in subsequent processes. As a consequence, less trustworthy sources, less informative, or implausible anchors will lead to less selective activation of one's knowledge base and, thus, exert a smaller anchoring effect or even none at all.

Another question is how subjective plausibility is determined in the first place (before any subsequent retrieval cues are changed). Most probably, the anchor will be almost automatically compared to one's item-specific knowledge base. The result of this comparison process might be that the anchor either falls inside or outside the subjective range of plausible values (cf. the inclusion-exclusion model of Schwarz & Bless, 1992). The anchor will be more likely considered plausible in the first case than in the second. This decision, however, need not be dichotomous. If the distance from the center of one's knowledge base is taken into account, continuous degrees of plausibility or implausibility should be predictable.

In sum, the reported findings support the assumption of an evaluative stage that determines the subjective plausibility of an anchor value. The plausibility judgment could be based on the numerical extremity of the anchor (as compared to one's own knowledge base), the credibility of its source and its informative value. Depending on the outcome of this evaluation, that is, the degree of plausibility, an anchor will lead to more or less selective activation of one's knowledge base. As a consequence, implausible anchor values show less or even no anchoring, while plausible anchors lead to large anchoring effects. Individual differences in suggestibility, though, seemed to play no role in this process.

References

Burger, H. (1971). [*The suggestive influenceability of statements about observations - Development and first evaluation of a test on statement suggestibility (TAS) in relation to testing the credibility of children and juvenile witnesses*]. Unpublished Dissertation, Albert-Ludwigs-Universität Freiburg, Bamberg.

Campbell, J.D., & Tesser, A. (1983). Motivational interpretations of hindsight bias: An individual difference analysis. *Journal of Personality, 51*, 605-620.

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology, 66*, 460-473.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212-252). New York: Guilford.

Chapman, G. B., & Johnson, E. J. (1994). The limits of anchoring. *Journal of Behavioral Decision Making, 7*, 223-242.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Davies, M. F. (1992). Field dependence and hindsight bias: Cognitive restructuring and the generation of reasons. *Journal of Research in Personality, 26*, 58-74.

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin, 107*, 311-327.

Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin, 21*, 1161-1166.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*, 136-153.

Leary, M. R. (1982). Hindsight distortion and the 1980 presidential election. *Personality and Social Psychology Bulletin, 8*, 257-263.

Loftus, E. F. (1979). Reacting to blatantly contradictory information. *Memory and Cognition, 7*, 368-374.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion. Central and peripheral routes to attitude change.* New York: Springer.

Plous, S. (1989). Thinking the unthinkable: The effects of anchoring on likelihood estimates of nuclear war. *Journal of Applied Social Psychology, 19*, 67-91.

Pohl, R. F. (1995). *Quantifying the impact of single and double anchors in a hindsight-bias paradigm.* Paper presented at the SPUDM-15 (Subjective Probability, Utility, and Decision Making), Jerusalem, Israel.

Pohl, R. F. (1998). The effects of feedback source and plausibility on hindsight bias. *European Journal of Cognitive Psychology, 10*, 191-212.

Pohl, R. F. (2000). Hindsight bias: robust, but not reliable. Manuscript submitted for publication.

Pohl, R. F., & Eisenhauer, M. (1997). SARA: An associative model for anchoring and hindsight bias. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (p. 1103). Mahwah, NJ: Erlbaum.

Pohl, R. F., & Gawlik, B. (1995). Hindsight bias and misinformation effect: Separating blended recollections from other recollection types. *Memory, 3*, 21-55.

Pohl, R. F., & Hell, W. (1996). No reduction of hindsight bias with complete information and repeated testing. *Organizational Behavior and Human Decision Processes, 67*, 49-58.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 14, pp. 207-262). San Diego, CA: Academic Press.

Read, J. D., & Bruce, D. (1984). On the external validity of questioning effects in eyewitness testimony. *International Review of Applied Psychology, 33*, 33-49.

Rhodes, N., & Wood, W. (1992). Self-esteem and intelligence affect influenceability: The mediating role of message reception. *Psychological Bulletin, 111*, 156-171.

Rokeach, M. (1960). *The open and closed mind.* New York: Basic Books.

Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: Assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgment* (pp. 217-245). Hillsdale, NJ: Erlbaum.

Shiffrin, R. M., & Raaijmakers, J. G. W. (1992). The SAM retrieval model: A retrospective and a prospective. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in Honor of William K. Estes* (Vol. 2, pp. 69-86). Hillsdale, NJ: Erlbaum.

Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 544-551.

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology, 73*, 437-446.

Switzer, F. S., III, & Sniezek, J. A. (1991). Judgment processes in motivation: Anchoring and

adjustment effects on judgment and behavior. *Organizational Behavior and Human Decision Processes, 49*, 208-229.

Toglia, M. P., Ross, D. F., Ceci, S. J., & Hembrooke, H. (1992). The suggestibility of children's memory: A cognitive and social-psychological interpretation. In M. L. Howe, C. J. Brainerd, & V. F. Reyna (Eds.), *The development of long-term retention* (pp. 217-241). New York: Springer.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases, *Science, 185*, 1124-1131.

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General, 125*, 387-402.

The reported research was supported by grant Po 315/6-2 from the Deutsche Forschungsgemeinschaft (Bonn/Germany) and was carried out while I was working at the Catholic University of Eichstätt (Germany). I would like to thank Stephanie Lösch for her help in data collection and the members of the "Cognitive Illusion Research Group" for their invaluable support in theoretical as well as practical questions. Gerd Fessler provided very helpful comments on an earlier version of this paper.