Explaining the Enigmatic Anchoring Effect: Mechanisms of Selective Accessibility

Fritz Strack and Thomas Mussweiler
Universität Würzburg

Results of 3 studies support the notion that anchoring is a special case of semantic priming; specifically, information that is activated to solve a comparative anchoring task will subsequently be more accessible when participants make absolute judgments. By using the logic of priming research, in Study 1 the authors showed that the strength of the anchor effect depends on the applicability of activated information. Study 2 revealed a contrast effect when the activated information was not representative for the absolute judgment and the targets of the 2 judgment tasks were sufficiently different. Study 3 demonstrated that generating absolute judgments requires more time when comparative judgments include an implausible anchor and can therefore be made without relevant target information that would otherwise be accessible.

In current psychological research, few phenomena are easier to demonstrate and harder to explain than the so-called *anchoring effect*, a biased estimate toward an arbitrary value considered by judges before making a numerical estimate (Jacowitz & Kahneman, 1995). In one of the best known and most typical anchoring studies[1] (Tversky & Kahneman, 1974), research participants were given an arbitrary number between 0 and 100 and were asked to indicate whether the percentage of African nations in the United Nations was higher or lower than that number. Participants then estimated the actual percentage. Results clearly indicated that the final assessments were influenced by the initial value provided by the experimenter; participants who had received a relatively high number as a standard for the comparative judgments gave higher absolute estimates than participants who were given a lower number as a standard of comparison.

This finding, that an absolute numerical judgment may be assimilated toward the standard of a preceding comparative judgment, has been frequently replicated by using a diverse range of stimulus materials (e.g., Cervone & Peake, 1986; Jacowitz & Kahneman, 1995; Northcraft & Neale, 1987; Plous, 1989; Switzer & Sniezek, 1991). Moreover, anchoring seems to be related to other judgmental phenomena, like overattribution (Leyens, Yzerbyt, & Corneille, 1996; Quattrone, 1982) and hindsight (Pohl & Hell, 1996). Although the described anchoring effect is considered to be strong, robust, and reliable, little agreement exists about the psychological mechanism underlying this ubiquitous phenomenon.

To account for the anchoring phenomenon, Tversky and Kahneman (1974) originally suggested a mechanism in which the anchor serves as a starting point for adjustment. In a more detailed elaboration of the argument, Jacowitz and Kahneman (1995) assumed that judges who are first asked to determine whether a target value is higher or lower than that of a given

The present research was supported by a grant from the Deutsche Forschungsgemeinschaft. We thank Rüdiger Pohl for providing methodological advice and calibration data. We are grateful to Marti Hope Gonzales for valuable suggestions concerning this article and to Klaus Fiedler, Jens Förster, Bettina Hannover, Tory Higgins, Roland Neumann, Norbert Schwarz, Sabine Stepper, Lioba Werth, and Werner Wippich for helpful discussions of the issues. Thanks also go to Martin Baierl, Wolfgang Danneil, Uta Deppe, Gunther Fritze, Christian Harneit, Klaus Hümpfner, Marcus Neumann, Anke Siebers, Anne Sequeira, Réné Thyrian, and Lioba Werth for helping to collect the data and to Jens Unger for the computer programming.

Correspondence concerning this article should be addressed to Fritz Strack or Thomas Mussweiler, Psychologie II, Universität Würzburg, Röntgenring 10, D-97070 Würzburg, Germany. Electronic mail may be sent via Internet to strack@psychologie.uni-wuerzburg.de or to mussweiler@psychologie.uni-wuerzburg.de.

[1] Judgmental assimilation toward an arbitrary number has also been obtained by using other paradigms. For example, Tversky and Kahneman (1974) reported that the result of calculating a product was greater if the computation started with the highest rather than with the lowest number. Wilson, Houston, Etling, and Brekke (1996) found anchoring effects when participants copied five pages of numbers. Kruglanski and Freund (1982) demonstrated anchoring effects in probability estimates of conjunctive and disjunctive events. On the surface, these situations resemble the standard anchoring paradigm; psychologically, however, they may be based on different mechanisms. Analogously, judging the prevalence of seven-letter words of the form "____n_" to be less likely than words of the form "____ing" and believing that Linda is less likely to be a bank teller than a bank teller and an active feminist are both instances of the *conjunction effect* (Tversky & Kahneman, 1983). However, the same conjunction effect may be produced by different judgmental mechanisms, namely availability in the first case and representativeness in the second. From this perspective, anchoring seems to be less a heuristic and more an effect, like the conjunction fallacy.

Journal of Personality and Social Psychology, 1997, Vol. 73, No. 3, 437–446
 0022-3514/97/$3.00

anchor adjust their estimates in the appropriate direction until an acceptable value is found. Because this adjustment process terminates at the nearest upper or lower boundary of a large range of acceptable values, it is considered to be generally insufficient.

Note that this explanation presupposes that the given anchor is outside of a range of acceptable values. That is, an adjustment to the boundary of that range is possible only if the given anchor is more extreme than the boundary value itself. However, anchoring effects do not depend on this restriction. Rather, such influences have frequently been observed when the anchors seemed to be as plausible as the target value. Northcraft and Neale (1987), for example, found anchoring effects by using prices of $65,900, $71,900, $77,900, and $83,900 as anchors for estimating the cost of a house with an actual listing price of $74,900. Because all of these anchors constituted plausible and acceptable values, judges did not know how far they should adjust their response. Thus, it is difficult for the described adjustment process to account for the obtained anchoring effect. Therefore, an alternative mechanism might be fruitfully invoked for situations in which the anchor is a plausible candidate for the target value.

The Anchoring Paradigm and Its Judgmental Strategies

To account for the broad array of anchoring effects, it seems necessary to take a closer look at the basic experimental paradigm, which consists of two separate judgment tasks: a comparative and an absolute judgment. Typically, the experimental manipulation provides participants with a standard of comparison. To form the judgment, participants must retrieve or generate information that is relevant to the task at hand. This may be achieved in at least three ways and depends on what the judges know about the target.[2] This knowledge determines the range of values that are considered plausible: The less the person knows, the wider the range.

First, a person may know the true value of the target. If a judge believes that a certain number is the actual value, this knowledge may be used to answer the comparative question. For example, if a person knows (or believes) that the Mississippi River is exactly 2,350 miles long, this number can be easily compared to any standard value, be it 25,000 or 2,000 miles.

Alternatively, an idealized judge may have no knowledge about the individual target but may possess some generic knowledge about a category to which the target belongs. For example, a person may know that rivers in general have certain minimal and maximal extensions and may use this knowledge to solve the comparative task. Thus judges who merely know that the Mississippi is a river may apply their subjective range of plausible lengths of rivers and decide whether the target is above the upper or below the lower boundary of this range. If such a person is asked whether the Mississippi is longer or shorter than 25,000 miles, her or his generic knowledge about rivers would suffice to generate the comparative answer. It is important to note that categorical knowledge is useful only when the anchor value lies outside of the category boundaries. Had 2,000 miles been provided as a standard of comparison, categorical knowledge about rivers in general would have been less relevant.

However, the described mechanisms do not apply to many anchoring tasks. That is, judges rarely know the exact target value, and a simple categorical judgment does not suffice, because anchor values are typically less extreme than the category boundaries.[3] To solve the task in these situations, judges may have to engage in more complex cognitive operations and form a mental model (Johnson-Laird, 1983) in which the target may be conceptually and imaginally represented and tested against the anchor value. To perform such a test, participants are likely to entertain the possibility of the target possessing the anchor value and create a corresponding mental representation. Klayman and Ha (1987) considered such a *positive test strategy* a rational all-purpose procedure for hypothesis testing, and Chapman and Johnson (1994) suggested that this strategy plays a role in anchoring tasks. For example, if the question asks whether the Mississippi River is longer or shorter than 3,000 miles, judges may imagine the north–south extension of the United States and use their geographic knowledge to compute the answer. In contrast, if the anchor value is 200 miles, it is sufficient to consider the extension of the state of Louisiana.

In sum, the three strategies differ with respect to their knowledge base and have different implications for the subsequent absolute judgment, which captures the anchoring effect. In the first case, when participants know what they believe to be the true value, the absolute estimate should not be affected by the previous comparative judgment. If participants know that the Mississippi River is 2,350 miles long, they are likely to use this knowledge for both judgment tasks, just as a person's honest report of her age would be uninfluenced by her previous answer to a question asking if she was older or younger than 40. Curiously, this possibility has never been investigated, probably because it does not describe a situation of uncertainty, which is seen as a precondition of judgmental heuristics (e.g., Tversky & Kahneman, 1974).

When an anchor value is implausible and generic categorical knowledge is the only basis for answering the comparative question, little information is provided about the actual value of the target. In this case, anchoring may occur in the form of a judgmental adjustment that will terminate at the nearest boundary of the range of acceptable values. Judges may then either use the boundary value or further test the possibility that the boundary value is true.

For judges who generate a mental model because they have to compare the target with a plausible standard, we expect that the information activated to construct such a mental model will be more accessible when the absolute judgment is formed. If participants pursue a positive test strategy, anchor-consistent information will be more likely to be recalled. To account for the effects of anchors that lie within the boundaries of the global category, we propose that participants construct a mental model that selectively increases the accessibility of anchor-consistent information. This information will be more likely to be used when the subsequent absolute question is answered.

---

[2] These strategies are ideal types and may operate jointly in many situations.

[3] To be sure, judges may find narrower subcategories for which the anchor values are beyond boundaries. For example, some judges may have a category of *large rivers* for which a moderate anchor may be beyond the lower boundary.

This notion has interesting implications: If anchoring is due to selectively increased accessibility of anchor-consistent information, it should be influenced by those factors that have been found to determine the use of such accessible information. These determinants were integrated into a judgmental model (Strack, 1992) that deems two characteristics of accessible information crucial: its applicability and its representativeness.

The first precondition for information use is considered to be its applicability. As previous research has demonstrated (e.g., Higgins, Rholes, & Jones, 1977; Banaji, Hardin, & Rothman, 1993; Higgins & Brendl, 1995), activated information that does not apply to the characteristics of a particular target is not used for its categorization. In the context of the anchoring paradigm, this implies that information that is accessible but inapplicable to the feature to be judged will be less likely to be used to generate an assessment, even if it contains a potential response. Study 1 was designed to test this hypothesis.

Second, the information must be representative. It has been argued (Strack, 1992) that the way information is used depends on its representativeness with respect to the target. That is, if activated content is closely related to the target, it will be used as a basis for judgment and will yield an assimilation effect; if it is not considered representative, an activated content may either be excluded from a judgment (Martin, 1986; Schwarz & Bless, 1992) or used as a standard of comparison, which leads to contrast (for examples, see Herr, 1986; Strack, Schwarz, & Gschneidinger, 1985). Applied to the anchoring paradigm, this analysis suggests that applicable information may lead to a contrast effect if it is not representative. This implication was tested in Study 2.

The third study tested implications of the accessibility notion for situations in which the comparative judgments are not assumed to involve the activation of relevant information. Specifically, we tested predictions about the different cognitive mechanisms for plausible and implausible anchor values.

Study 1

To examine the accessibility of retrieved and generated information as a crucial determinant of the anchoring effect, we manipulated the anchoring dimension in the first study and allowed the targets of the comparative and the absolute judgments to remain the same. This variation also served as a test of the numerical-prime explanation, which assumes that anchoring occurs because the anchor itself primes a candidate answer (see Jacowitz & Kahneman, 1995). Were this the case, a change of the response dimension would not matter, and the same assimilation effects would be observed. However, if the anchoring effect is mediated by higher accessibility of information that is generated to solve the initial comparative task, the applicability of this information becomes crucially important. As a consequence, the dimension on which the absolute judgment is provided plays a critical role.

For example, if the anchoring task asks judges to compare the length of the Mississippi River, the value of the anchor should exert a stronger effect when the subsequent target judgment also refers to its length than when it refers to its width. This is because the information generated to assess an object on one dimension is less applicable when the same object is judged on a different dimension.

Study 1 was conducted to test the first implication of the selective accessibility model, namely that the activated information must be applicable to the target judgment. Because all studies were similar in procedure, Study 1 is described in detail, whereas, for the remaining studies, we note procedural deviations when they occur.

Method

Participants. We recruited 32 male and female nonpsychology students at the University of Würzburg as participants. They were asked to participate in a pretest for the construction of a questionnaire assessing general knowledge and were offered a chocolate bar as compensation. The participants were randomly assigned to one of four experimental conditions.

Materials. The questionnaire consisted of one pair of questions. Congruent with the typical anchoring procedure, the first question asked for a comparative judgment and the second question for an absolute judgment. Thus participants were first asked to indicate whether a quantitative value of a target was higher or lower than a particular anchor value; they were then asked to give an exact estimate of that quantitative value. For example, they were first asked whether the Brandenburg Gate in Berlin is taller or shorter than 150 m and were then asked how tall the Brandenburg Gate is.

The anchors were chosen such that they differed with respect to the direction in which they deviated from the mean of the answers of a calibration group ($n = 28$) that answered absolute questions only. More specifically, they were set at the 15th and 85th percentile of the calibration answers. Half of the participants received a high anchor, and half received a low anchor.

The first and the second questions always pertained to the same object, but we varied the dimension to which they referred. For half of the participants, both questions referred to the same dimension; for the other half, they referred to two different dimensions. For example, after being asked whether the Brandenburg Gate is taller or shorter than 150 m, participants were asked either how high or how wide the Brandenburg Gate is. To control for content effects, half of the participants received questions pertaining to the Brandenburg Gate, and the other half received questions pertaining to the Cathedral of Cologne. The anchors used for these two objects are listed in Table 1. To control for effects of the dimension, both dimensions were used in the first and the second questions for half of the participants. Thus the four experimental conditions resulted from a combination of high versus low anchors and change versus no change of judgmental dimension.

Procedure. Participants were recruited in the university cafeteria to participate in a pretest for the construction of a general-knowledge questionnaire and received a chocolate bar for their participation. Participants completed the questionnaire in groups of up to 15.

The instructions emphasized that the questionnaire was meant to find the best wording for general-knowledge questions. To reduce the ascribed informativeness of the anchors given, the instructions pointed out that the values were randomly selected both to minimize their impact on the answers and to identify the impact of different question formats. Participants were also asked to answer all the questions in order of appearance.

Results

To pool answers across different content dimensions,[4] participants' responses were transformed into z scores for each ques-

[4] A preliminary analysis revealed that the specific content exerted no effect ($F < 1$) for the main effect and the interaction effects involving content domain.

Table 1
Objects and Anchors Used in Study 1

Object of first question	Object of second question: Same dimension	Object of second question: Different dimension	Anchors: High	Anchors: Low
Height of the Brandenburg Gate	Height of Brandenburg Gate	Width of the Brandenburg Gate	150	25
Width of the Brandenburg Gate	Width of the Brandenburg Gate	Height of the Brandenburg Gate	150	25
Length of the Cathedral of Cologne	Length of the Cathedral of Cologne	Height of the Cathedral of Cologne	320	60
Height of the Cathedral of Cologne	Height of the Cathedral of Cologne	Length of the Cathedral of Cologne	320	60

Note. Values for high and low anchors are in meters.

tion.[5] Thus the resulting cell means reflect participants' average deviations from the question mean in units of the pertinent standard deviation.

As is apparent from Table 2, the typical anchoring effect was replicated such that a high anchor value led to higher estimates than did a low anchor value. More interesting, this assimilation effect was much stronger when the absolute extension judgment was provided on the same dimension as the preceding comparative judgment; the effect was weaker when the two judgments were given on different dimensions of longitudinal extension.

This pattern of means was borne out in a 2 (high vs. low anchor) × 2 (same vs. different dimension) between-subjects analysis of variance (ANOVA) with the *z*-transformed answers to the second question as dependent variables. In this analysis, the main effect for the anchoring manipulation was statistically significant, $F(1, 28) = 18.23, p < .001$; more important, however, this main effect was qualified by a significant interaction effect, $F(1, 28) = 10.50, p < .005$, indicating the reliability of the finding that the anchor effect was stronger when the judgmental dimensions were the same than when they were different. The significant main effect for same versus different dimension, $F(1, 28) = 7.76, p < .01$, reflects that the difference in the size of anchoring for the two dimension conditions was mainly due to the stronger effect of high anchors in the same dimension condition.

Discussion

This finding of a diminishing assimilation effect for anchor values associated with a dimension other than the target dimension cannot be accounted for by numeric priming. That is, the strength of the observed assimilation effect is not exclusively due to the mere activation of a potential response. Hence, these results call for an explanation that considers aspects of the anchoring task other than the anchor value itself. Specifically, the present findings suggest that the strength of the anchoring effect depends on how applicable the activated information is perceived to be. Information about the height of the Brandenburg Gate is more applicable when its height is judged than when its width is assessed. Similarly, information about the height of the Cathedral of Cologne is more applicable when judging its height than when judging its length. In standard anchoring tasks, the dimensions of the comparative and the absolute judgment are identical. Therefore, we cannot determine whether the anchoring effect was due to the activation of the numerical anchor value or to the activation of the information that was used to compare the target with the anchor value. The present findings suggest that the latter is the case.

A second strategy to find out whether the anchoring effect is caused by the information that is activated in the comparative task is to study the direction of the influence. Whereas a numerical-priming explanation would predict anchoring effects to produce only judgmental assimilation toward the anchor, an explanation that focuses on the activated information may also account for contrast effects. It has been argued (e.g., Martin, 1986; Strack & Martin, 1987) that to predict the direction of priming effects one must not only consider whether information is activated but must also consider how that information is used. For example, if information is not representative (Strack, 1992) for the judgment at hand, it may be excluded (Schwarz & Bless, 1992) from the target category and used as a standard of comparison. As Schwarz and Bless have shown, such an exclusion may lead to a contrast effect and change the judgment in the opposite direction of the implication of the prime. Applied to the anchoring task, a change of the object may cause the information that was activated to solve the comparative task to be ignored as a basis for the absolute judgment and instead cause it to be used as a standard of comparison. However, the likelihood of a contrast effect should be higher the more dissimilar the target and context stimuli are to one another (see Herr, 1986; Herr, Sherman, & Fazio, 1983; Strack et al., 1985). Herr et al.

Table 2
Overall z Values, Study 1

Anchor	Dimension: Same	Dimension: Different
High	1.20	−0.20
Low	−0.55	−0.44

Note. $n = 8$ for all cells.

[5] A comparison across content domains must not only cope with different mean values but must also provide for different variances. The means are mainly determined by the characteristics of the targets, whereas the variance is a function of judges' knowledge about the target. Thus, in absolute values, apparent differences in the strength of anchoring effects may be greatly misleading when reported as unstandardized values. Therefore, tables listing the standardized values are included in the text, whereas the absolute values can be found in the Appendix.

(1983) found that when the ferocity of an animal had to be judged, priming of a moderately ferocious exemplar would produce an assimilation effect, whereas priming an extreme exemplar would lead to contrast. Strack et al. (1985) found that describing a positive or a negative life event would produce an assimilation on judgments of well-being when the event happened in the recent past and a contrast effect when it happened in the distant past. In sum, these results suggest that objects that are dissimilar to the target may lead to contrast effects in judgments of the target. Applying this finding to the anchoring paradigm, we assumed that performing the comparative anchoring task with an object that is dissimilar to the target object would lead to a contrast effect.

Study 2

Method

Participants. We recruited 32 male and female nonpsychology students at the University of Trier as participants. As in the previous study, they were asked to participate in a pretest for the construction of a questionnaire assessing general knowledge and were offered a chocolate bar as compensation. The participants were randomly assigned to one of four experimental conditions.

Materials. The questionnaire consisted of two pairs of questions similar to those described in Study 1. The first and the second question pertained either to the same stimulus or to a stimulus that differed greatly on the relevant dimension. For example, after indicating whether the mean winter temperature in the Antarctic is higher or lower than −50°C, participants were asked how high the mean winter temperature is in the Antarctic or in Hawaii.

Procedure. The high and low anchors were determined by selecting values that were about one standard deviation above or below the mean of an independent calibration group (n = 151); stimuli and anchors are listed in Table 3.

To control for content effects, each of the four conditions (high vs. low anchor, same vs. different object) was realized with two questions pertaining to different contents; in addition, we counterbalanced the order of the questions. Thus two different questionnaires were administered in each condition. The procedure was the same as in Study 1.

Results

Inspection of Table 4 reveals that the direction in which high versus low anchors influenced the judgments depended on whether target and context stimuli were identical. When target and context stimuli were identical, the previous anchoring effect was replicated; participants' judgments were assimilated toward the anchors. However, when context and target stimuli were not identical, anchoring led to contrast. A 2 (high vs. low anchor) × 2 (same vs. different object) between-subjects ANOVA with the z-transformed answers for the target questions as dependent variables yielded a significant interaction effect, $F(1, 28) = 6.30$, $p < .02$, whereas the main effect for the anchor value was nonsignificant, $F(1, 28) < 1$. Note that the same pattern appears for both questions, although separate analyses revealed the effect to be stronger for the Antarctic question.

Table 3
Objects and Anchors Used in Study 2

Object of first question	Object of second question: Same	Object of second question: Different	Anchors: High	Anchors: Low
Aristotle	Aristotle	Kant	−90[a]	−600[a]
Antarctic	Antarctic	Hawaii	−20[b]	−50[b]

[a] Year of birth. [b] °C.

Table 4
Overall z Values, Study 2

Anchor	Object: Same	Object: Different
High	0.22	−0.34
Low	−0.22	0.34

Note. n = 8 for all cells.

Discussion

The results of Study 2 show that anchors may exert directionally different effects depending on whether the object of the comparative task is identical with or distinctly different from the object of the absolute judgment. Although the conventional assimilation effect of anchoring was replicated when the two objects were identical, a contrast effect was obtained if they were different. Under those conditions, a high standard of comparison led to a lower absolute judgment, and vice versa.

This result provides further evidence that the anchoring effect is not due to the mere activation of a numerical value but to the information that is activated in the comparative task. In line with previous findings on the divergent effects of priming (e.g., Herr, 1986; Strack, Martin, & Schwarz, 1988), information that is accessible for a judgment may be differentially used. One determinant, the identity versus dissimilarity between a target and a context stimulus, was manipulated in the present study and found to determine the direction of an anchor's influence. Thus the present outcomes parallel existing findings about the activation and use of information and suggest that similar mechanisms may be involved in judgmental anchoring.

A third consequence of semantic priming is that response latencies are facilitated if semantically related information is presented prior to the target (for a review, see Neely, 1991). For example, Neely (1977) demonstrated that, in a lexical decision task, response latencies for the word *robin* were shorter when the category label *bird* was presented beforehand. In a related vein, Fazio, Sanbonmatsu, Powell, and Kardes (1986) found that the evaluation of positive and negative adjectives was facilitated by the prior presentation of attitude objects with the same valence. Applied to the anchoring paradigm, this mechanism implies that solving the comparative task facilitated the absolute judgment to the extent that the comparative answer was based on relevant information, that is, when judges engaged in an elaborate test strategy and formed a mental model of the target object. As we have previously argued, this should be the case when the anchor lies within the range of plausible values

but not when the anchor value is implausible. Thus, if a person has to decide whether the Mississippi River is longer or shorter than 25,000 miles, the comparative task can simply be solved by applying categorical knowledge. Knowing that no river reaches this extension, the person need not form an elaborate mental model of the target to arrive at an appropriate answer. At the same time, however, little applicable information will be accessible to facilitate the absolute judgment.

This reasoning has four testable consequences. First, because the possibility of a boundary value is considered for the absolute judgment, an implausible anchor that lies outside the range of possible values for the target category (e.g., 25,000 miles for the category *rivers*) should produce a greater assimilation effect than a plausible anchor. Second, because no elaborate test strategy needs to be performed, the comparative task should be solved faster when the standard is implausible than when it is plausible. Third, because an elaborate test strategy needs to be performed for the comparative task, plausible anchors activate applicable information and facilitate the absolute judgment. Therefore plausible anchors should lead to shorter latencies for the absolute judgments. Finally, the more elaborate the comparative test, the more facilitation will result for the absolute judgment. Therefore latencies for the comparative and the absolute judgments should be negatively correlated. However, this should only be the case when the increased time that is spent on the comparative judgment is likely to activate information applicable to the absolute judgment. As a consequence, negative correlations for the response latencies should be obtained only for plausible anchors. These predictions were tested in the third study.

Study 3

Method

Participants. We recruited 69 students at the University of Würzburg as participants and asked them to participate in a pretest for the construction of a questionnaire assessing general knowledge. A chocolate bar was offered as compensation. Two participants were excluded from the analysis because of missing data.

Materials. The questions used were similar to those of Studies 1 and 2. The anchors were chosen such that they differed in both their direction and their plausibility. The latter was determined by asking 40 different participants to assess the plausibility of comparative questions by using plausible and implausible anchors on a 5-point rating scale ranging from 1 (*absolutely implausible*) to 5 (*very plausible*). Plausible anchors deviated about 1 standard deviation from the mean of the calibration group (n = 151); implausible anchors deviated from this mean by more than 10 standard deviations, except in instances in which such an extreme deviation yielded logical inconsistencies. In addition, for any anchor to qualify as plausible or implausible, more than 80% of the participants had to assign the potential anchor to one of the two extreme categories on the rating scale. As a result, four different types resulted from the orthogonal combination of plausibility (plausible vs. implausible) and direction (high vs. low). They are listed in Table 5.

To control for content and order with a Latin-Square rationale, each type of anchor was used in all question pairs, whereas the order of the questions was kept constant. To allow for an assessment of response latencies, the questions were presented on a personal computer.

Procedure. Participants took part in the experiment in groups of up to 4. They were recruited in the university cafeteria and were then led to a personal computer in an adjacent room. The same instructions used in the other experiments were displayed on the computer screen, and the experimenter demonstrated how to provide the answers by using the keyboard. Participants were instructed to answer the comparative question by pressing either the q key, which was marked with a red sticker, or the p key, which was marked with a green sticker. For each comparative question, the keys corresponding to the two possible answers (e.g., longer or shorter) were depicted on the bottom of the computer screen. To reduce variance in response latencies, participants were told to position their forefingers on the two keys before the question appeared on the screen. Participants were told to use the number pad on the keyboard to answer the absolute judgment questions. Finally, participants were instructed to answer the questions as accurately and as quickly as possible. They were then presented 22 pairs of comparative and absolute questions. The first 14 pairs served as practice trials to become acquainted with the experimental procedure. Pairs 15 through 22 were the critical trials and were thus included in the analysis.

Before each question, a focus point appeared in the center of the screen for .4 s, followed by the question, which remained on the screen until the first answer key was pressed. After a pause of 3 s, the next question was presented in the same sequence.

Results

Anchoring effect. Table 6 reveals that a substantial anchoring effect again was found: Overall, high anchors led to higher absolute judgments than did low anchors. This difference yielded a significant main effect of direction in a 2 (high vs. low anchors) × 2 (plausible vs. implausible) within-subjects multivariate analysis of variance (MANOVA) with the standardized answers to the eight critical absolute questions as dependent

Table 5
Objects and Anchors Used in Study 3

Question	Actual value	Plausible anchors		Implausible anchors	
		High	Low	High	Low
Antarctic: mean temperature in winter (°C)	−68	−17	−43	45	−210
Einstein: year of first visit to United States	1921	1939	1905	1992	1215
Da Vinci: year of birth	1452	1698	1391	1952	−300
Gandhi: age	78	79	64	140	9
Ulm: altitude (m)	478	320	150	10,500	−1,700
Aristotle: year of birth	−322	−220	−490	1832	−25,000
Whale: length (m)	33.0	49.0	21.0	900.0	0.2
Elbe: length (km)	1,165	890	550	45,000	25

Table 6
Overall z Values, Study 3

Anchor	Plausible	Implausible
High	.04	.25
Low	−.13	−.17

Note. $n = 67$ for all cells.

variables, $F(1, 66) = 7.61$, $p < .01$. Moreover, plausibility yielded no significant effect: Anchoring occurred for plausible as well as implausible anchors, $F(1, 66) = 3.55$, $p < .07$, for the interaction and $F(1, 66) < 1$, for the main effect. A mere inspection of the means reveals that implausible anchors were at least as effective as plausible ones.

Response latencies. More interesting are the response latencies for the comparative and absolute answers. As suggested by Fazio (1990), we conducted logarithmic transformations of response latencies to reduce the skewness of the response distribution. Inspection of Table 7 reveals that for the comparative task, response latencies were longer for plausible than for implausible anchors, whereas for the absolute task this pattern was reversed. Here, plausible anchors yielded shorter response latencies than implausible anchors. The corresponding two-way Plausibility × Response Type interaction proved to be significant in a 2 (high vs. low anchors) × 2 (plausible vs. implausible anchors) × 2 (comparative vs. absolute responses) MANOVA with transformed response latencies for the answers to the eight critical comparative and absolute questions as dependent variables, $F(1, 66) = 72.81$, $p < .001$. Moreover, plausible anchors yielded shorter response latencies than implausible anchors, $F(1, 66) = 23.78$, $p < .001$, and absolute questions yielded shorter response latencies than comparative questions, $F(1, 66) = 38.81$, $p < .001$. Response latencies did not differ for high and low anchors: $F(1, 66) = 1.35$, $p < .25$, for the main effect of direction; $F(1, 66) = 1.88$, $p < .2$, for the two-way Direction × Plausibility interaction; $F(1, 66) = 2.55$, $p < .12$, for the two-way Direction × Response Type interaction; and $F(1, 66) < 1$, for the three-way interaction.

Separate analyses for comparative and absolute questions revealed that the main effect of plausibility in a 2 (high vs. low anchors) × 2 (plausible vs. implausible anchors) MANOVA was significant for both response types: For the comparative task, $F(1, 66) = 8.46$, $p < .005$, and for the absolute task, $F(1, 66) = 61.31$, $p < .001$. Again, the response latencies did not differ for high versus low anchors: Neither the corresponding main effects, $F(1, 66) < 1$, for the comparative task and $F(1, 66) = 2.62$, $p < .11$, for the absolute task nor the interaction effects, $F(1, 66) < 1$, for the comparative task and $F(1, 66) = 1.70$, $p < .20$ for the absolute task, attained significance.

These results suggest that, for plausible anchors, participants solve the comparative task by engaging in an elaborate and time-consuming test. Subsequently, however, relevant information is easily accessible and accelerates responses to the absolute question. If this logic is correct, it follows that the more time judges spend on the comparative task, the less time they need for answering the absolute question. Thus the response latencies should be negatively correlated. However, this should be the case particularly if the comparative task is solved by using information that is relevant for answering the absolute question, that is, when the target has to be compared with a plausible anchor.

Results from a correlational analysis support this prediction. For plausible anchors, comparative and absolute response latencies were negatively correlated ($r = -.21$, $p < .05$) such that longer response latencies for the comparative task were associated with shorter response latencies for the absolute task. In contrast, for implausible anchors, the response latencies for the two tasks were virtually unrelated ($r = -.02$, *ns*).

Discussion

The results of Study 3 provide further support for the hypothesis that mechanisms of semantic priming may be responsible for anchoring effects. Specifically, we found that the comparative task took more time when its solution was assumed to require an elaborate test (i.e., for plausible anchors) than when it was assumed to be solvable on the basis of categorical knowledge (i.e., for implausible anchors). In line with predictions and findings from priming research, answering the comparative question in an elaborative fashion shortened the latency of the absolute questions, presumably because relevant information was activated and subsequently more accessible. This was observed not only when the plausibility of the anchors was manipulated but also when the correlations of the response latencies between the comparative and the absolute task were analyzed. The more likely the activation of relevant information, the shorter the latency for the absolute question.

The results of Study 3 also speak to the question of whether anchoring in the implausible condition occurs as a simple adjustment to the boundary of the plausibility range. Perhaps judges simply select the first plausible value of their subjective distribution. The latency data suggest that this might not be the case. The fact that finding the absolute answer took more time after comparing the target with an implausible than with a plausible anchor suggests that judges may not have simply selected a boundary value but instead engaged in a more elaborative test. At this point, the data are merely suggestive, and more research about this issue is needed.

Finally, the fact that assimilation occurs with entirely implausible anchors (for similar results, see Chapman & Johnson, 1994; Quattrone et al., 1984, as cited in Plous, 1993) suggests that conversational influences are not necessary to account for anchoring effects; in the present study, the effect occurred with anchors that can hardly be interpreted as suggestive, such as Einstein visiting the United States in the year 1212 or the Ger-

Table 7
Logarithmic Transformation of Response Latencies

	Comparative question		Absolute question	
Anchor	Plausible	Implausible	Plausible	Implausible
High	8.87	8.79	8.41	8.86
Low	8.89	8.78	8.38	8.73

Note. $n = 67$ for all cells.

man river Elbe being 45,000 km long. In combination with findings from studies in which anchors were randomly chosen (e.g., Tversky & Kahneman, 1974) and with findings from studies that used elaborate cover stories, like the present research, the demonstrated impact of implausible anchors makes it difficult to reduce anchoring effects to a conversational phenomenon (see also Wilson et al., 1996).

General Discussion

Taken together, the present studies suggest that effects of an anchoring task in which participants have to compare a target with a plausible standard of comparison cannot be sufficiently understood when one focuses merely on the numerical value of the anchor. Rather it seems necessary to direct attention toward those cognitive mechanisms used to solve the comparative task and to the consequences of these mechanisms for generating absolute judgments. To account for the results, we propose a *selective accessibility model* that is based on fundamental findings about priming and the formation of judgments. Specifically, we suggest that, to solve the comparative task, participants test the possibility that the target possesses the anchor value and try to construct a mental model that includes information that is maximally consistent with the anchor value. Research from different areas (e.g., Higgins, 1996; Neely, 1991; Slamecka & Graf, 1978) suggests that such self-generated information will subsequently be highly accessible and provide a basis for relevant judgments (e.g., Ross, Lepper, Strack, & Steinmetz, 1977; Strack et al., 1985).

However, the mere accessibility of information is not sufficient to guarantee its use and determine the direction of its influence. First, to be used, activated information must be applicable or relevant to the judgment at hand (Higgins et al., 1977; Banaji et al., 1993; Higgins & Brendl, 1995). Although anchor values were identical in our studies, the results show that the comparative task only affected the absolute judgment when the activated information was applicable to the absolute judgment. Second, the direction of the influence depends on the representativeness of the accessible information (Strack, 1992) for the target judgment. That is, when the activated information is not representative, it may be excluded from the judgment (Martin, 1986; Schwarz & Bless, 1992), or it may be used as a standard of comparison and produce a contrast effect (for examples, see Herr, 1986; Strack et al., 1985). Thus both assimilation and contrast effects are possible manifestations of anchoring and must be studied with respect to the underlying cognitive mechanism.

Moreover, our findings show that activating relevant information not only influences the content of the absolute judgment but also the ease with which it is generated. That a response facilitation of the absolute judgment was found only for plausible, and not implausible, anchors suggests that the effect was not simply a function of comparing the target with a standard. Rather this finding speaks for the thesis that the increased accessibility of relevant information is the mechanism underlying anchoring effects.

It is worth noting that, within a selective accessibility model, the anchoring phenomenon is conceptually related to more general mechanisms of hypothesis testing (e.g., Bassok & Trope, 1984; Higgins & Bargh, 1987; Snyder & Swann, 1978; Trope & Bassok, 1982, 1983; Trope, Bassok, & Alon, 1984; Trope & Liberman, 1996). If the solution of the comparative task is akin to a test of the hypothesis that the target possesses the anchor value, insights about the cognitive mechanisms of hypothesis testing may be fruitfully applied to the anchoring situation. For instance, it can be expected that judges collect information that is relevant for a direct test of the hypothesis. As an example, participants who are asked to decide whether the Mississippi River is longer or shorter than 3,000 miles may act as if they tested the possibility that the Mississippi River is 3,000 miles long. Trope and his collaborators (Trope & Bassok, 1982, 1983; Bassok & Trope, 1984; Trope et al., 1984; Trope & Liberman, 1996) have argued that in such a situation participants rely primarily on diagnostic information that is consistent with the hypothesis but not with possible alternatives. In the context of the concrete anchoring example, information about short stretches of the Mississippi River, such as it extends from the northern to the southern border of the state of Louisiana, has little diagnostic value when one is trying to decide whether the Mississippi River is longer or shorter than 3,000 miles. Highly diagnostic is the information that the Mississippi extends from the northern to the southern border of the United States, because this feature of the target is closer to the hypothesis being tested. Therefore the latter information should be more likely to be activated in comparative tasks. As a consequence, this information will be highly accessible when the absolute judgment is generated at a later time and will provide the basis for a judgmental bias. Thus, the positive test strategy suggested by Klayman and Ha (1987) seems to generate information that improves the validity of comparative judgments but distorts subsequent judgments for which no specific hypothesis has been provided.

In various critiques of Snyder and Swann's (1978) seminal study, it has been noted that their evidence for a confirmation bias in hypothesis testing hinges on the fact that only one hypothesis was provided and that participants were not asked to consider an alternative possibility (Bassok & Trope, 1984; Semin & Strack, 1980; Trope & Bassok, 1982, 1983; Trope et al., 1984). On the surface, the anchoring task seems to consist of a test of two alternatives, namely that the characteristic of a target is either above or below a standard of comparison. Psychologically, however, judges appear to act as if they decided whether the target possesses exactly the value of the standard.

The present findings suggest that to prevent a confirmation bias it is necessary to consider alternative hypotheses. This, however, is not feasible in the context of a simple comparative judgment, in which the alternatives seem to be implied in the specific task. Providing two standards (and presumably two hypotheses) would transform the comparative task into a task in which the target has to be assigned to one of three ordinal positions, namely below both anchors, between the two anchors, or above both anchors. If selective accessibility is the mechanism that produces the anchoring effect, the described transformation of the comparison task should be an effective means of debiasing.

Finally, the selective accessibility mechanism explains why the anchoring bias is so robust in standard comparison tasks. Unlike in other judgmental heuristics, there exists no better, systematic way to solve the problem. Even under optimal cir-

cumstances, judges will use the most accessible evidence. Moreover, because the biasing information is not provided by an environmental source but is generated by judges themselves, participants may not be aware of the contaminating influence and may fail to engage in correction (see Strack, 1992; Strack, Schwarz, Bless, Kübler, & Wänke, 1993). Explicitly instructing participants to do so would require knowledge about the presumed influence (Petty & Wegener, 1993; Strack, 1992; Wegener & Petty, 1995). However, given the time and effort experimental psychologists have spent on solving the enigma of the psychological causes of anchoring, it seems unlikely that they were surpassed by the intuitive knowledge of their research participants.

References

Banaji, M. P., Hardin, C., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology, 65,* 272–281.

Bassok, M., & Trope, Y. (1984). People's strategies for testing hypotheses about another's personality: Confirmatory or diagnostic? *Social Cognition, 2,* 199–216.

Cervone, D., & Peake, P. K. (1986). Anchoring, efficacy, and action: The influence of judgmental heuristics on self-efficacy judgments and behavior. *Journal of Personality and Social Psychology, 50,* 492–501.

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

Fazio, R. H. (1990). A practical guide to the use of response latencies in social psychological research. In C. Hendrick & M. S. Clark (Eds.), *Research methods in personality and social psychology* (pp. 74–97). Newbury Park, CA: Sage.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50,* 229–238.

Herr, P. M. (1986). Consequences of priming: Judgment and behavior. *Journal of Personality and Social Psychology, 51,* 1106–1115.

Herr, P. M., Sherman, S. J., & Fazio, R. H. (1983). On the consequences of priming: Assimilation and contrast effects. *Journal of Experimental Social Psychology, 19,* 323–340.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Higgins, E. T., & Bargh, J. A. (1987). Social cognition and social perception. *Annual Review of Psychology, 38,* 369–425.

Higgins, E. T., & Brendl, M. (1995). Accessibility and applicability: Some "activation rules" influencing judgment. *Journal of Experimental Social Psychology, 31,* 218–243.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology, 13,* 141–154.

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

Johnson-Laird, P. N. (1983). *Mental models: Toward a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard University Press.

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

Kruglanski, A. W., & Freund, T. (1982). The freezing and unfreezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology, 19,* 448–468.

Leyens, J.-P., Yzerbyt, V., & Corneille, O. (1996). The role of applicability in the emergence of the overattribution bias. *Journal of Personality and Social Psychology, 70,* 219–229.

Martin, L. L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology, 51,* 493–504.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 3,* 226–254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–337). Hillsdale, NJ: Erlbaum.

Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes, 39,* 84–97.

Petty, R. E., & Wegener, D. T. (1993). Flexible correction processes in social judgment: Correcting for context-induced contrast. *Journal of Experimental Social Psychology, 29,* 137–165.

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

Plous, S. (1993). *The psychology of judgment and decision making.* New York: McGraw-Hill.

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

Quattrone, G. A. (1982). Overattribution and unit formation: When behavior engulfs the person. *Journal of Personality and Social Psychology, 42,* 593–607.

Ross, L. D., Lepper, M. R., Strack, F., & Steinmetz, J. (1977). Social explanation and social expectation: The effects of real and hypothetical explanations on subjective likelihood. *Journal of Personality and Social Psychology, 35,* 817–829.

Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 227–245). Hillsdale, NJ: Erlbaum.

Semin, G. R., & Strack, F. (1980). The plausibility of the implausible: A critique of Snyder and Swann (1978). *European Journal of Social Psychology, 10,* 379–388.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 592–604.

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

Strack, F. (1992). The different routes to social judgments: Experiential versus informational strategies. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgment* (pp. 249–275). Hillsdale, NJ: Erlbaum.

Strack, F., & Martin, L. L. (1987). Thinking, judging, and communicating: A process account of context effects in attitude surveys. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 123–148). New York: Springer.

Strack, F., Martin, L. L., & Schwarz, N. (1988). Priming and communication: Social determinants of information use in judgments of life satisfaction. *European Journal of Social Psychology, 18,* 429–442.

Strack, F., Schwarz, N., Bless, H., Kübler, A., & Wänke, M. (1993). Awareness of the influence as a determinant of assimilation versus contrast. *European Journal of Social Psychology, 23,* 53–62.

Strack, F., Schwarz, N., & Gschneidinger, E. (1985). Happiness and reminiscing: The role of time perspective, affect, and mode of thinking. *Journal of Personality and Social Psychology, 49,* 1460–1469.

Switzer, F. S., & 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.

Trope, Y., & Bassok, M. (1982). Confirmatory and diagnostic strategies in social information gathering. *Journal of Personality and Social Psychology, 43,* 22–34.

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

Trope, Y., Bassok, M., & Alon, E. (1984). The questions lay interviewers ask. *Journal of Personality, 52,* 90–106.

Trope, Y., & Liberman, A. (1996). Social hypothesis testing: Cognitive and motivational mechanisms. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 239–270). New York: Guilford Press.

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

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 91,* 293–315.

Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naive theories in corrections of perceived bias. *Journal of Personality and Social Psychology, 68,* 36–51.

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.

Appendix

Mean Values for Individual Questions Used in Studies 1 Through 3

	Same dimension		Different dimension	
Study/Question	High anchor	Low anchor	High anchor	Low anchor
Study 1				
Brandenburg Gate: height/width (m)	82.5	30.0	45.0	32.3
Cathedral of Cologne: height/length (m)	272.5	88.8	105.5	101.3
	Same object		Different object	
	High anchor	Low anchor	High anchor	Low anchor
Study 2				
Aristotle/Kant: year of birth	−247.5	−326.3	1735.5	1772.5
Antarctic/Hawaii: mean temperature in winter (°C)	−38.8	−51.5	22.3	23.3
	Plausible anchor		Implausible anchor	
	High anchor	Low anchor	High anchor	Low anchor
Study 3				
Antarctic: mean temperature in winter (°C)	−17.1	−36.6	−37.2	−59.8
Aristotle: year of birth	−297.9	−362.7	−139.5	−1150.4
Whale: length (m)	60.1	29.1	141.7	20.5
Da Vinci: year of birth	1543.6	1410.3	1632.9	1811.1
Ulm: altitude (m)	276.2	233.4	558.1	490.0
Einstein: year of first visit to U.S.	1807.3	1817.4	1928.2	1922.1
Elbe: length (km)	676.1	645.1	10124.6	284.7
Gandhi: age	67.9	99.6	66.7	50.1

Received March 27, 1995
Revision received December 16, 1996
Accepted December 16, 1996 ■