IGNORANT EXPERTS – HEIGHTENED CONFIDENCE UNDERMINES THE BENEFICIAL EFFECT OF A FOREWARNING ABOUT THE ANCHORING EFFECT

Inauguraldissertation

zur Erlangung des Doktorgrades

der Humanwissenschaftlichen Fakultät

der Universität zu Köln

nach der Promotionsordnung vom 10.05.2010

vorgelegt von

Ulrich Schlattmann

aus

Aachen

Mai 2012

Diese Dissertation wurde von der Humanwissenschaftlichen Fakultät der Universität zu Köln im Juli 2012 angenommen.

CONTENTS

ACKNOWLEDGMENTS

In the process of developing, conducting, and writing this dissertation, I received much help from many colleagues and friends. I would like to use this opportunity to thank all of them.

Special thanks go to Birte Englich, who made this dissertation possible in the first place. Her expertise, understanding, and patience added considerably to my graduate experience. I also thank Thomas Mussweiler a lot for the possibility to be a part of Social Cognition Cologne. In this vivid scientific research group I attended interesting talks, numerous interesting scientific debates and received important feedback.

It is a pleasure for me to thank my highly appreciated colleagues at Social Cognition Cologne, Barbara Stoberock, Susanne Schmittat, Janina Steinmetz, Ann-Christin Posten, Jennifer Mayer, Jan Crusius, Katja Corcoran, Femke van Horen, Gayanée Kedia, Pascal Burgmer, Mathias Forstmann, Tanja Hundhammer, Jens Lange, and Katharina Banscherus for their valuable comments and suggestions concerning my work, but also for the great atmosphere in this team. I am especially indebted to the interns Paula Viefhaus, Karim Ballach and Julia Knospe but also to Helmut Appel, Jasmin Schmitz, Miriam Schwarzenthal, and Claudia Beck who collected the data and helped me with the execution of my studies.

Finally, I would like to express my sincere gratitude to my parents and my friends − especially Alice von Welck and Kurt Michalczyk − who enrich my entire life and this dissertation with their ideas, their support and interest.

Thank You!

ABSTRACT

Anchoring effects are remarkably robust and difficult to correct. Forewarnings about the influence of anchor values have not led to a reduction of the anchoring effect (Wilson, Houston, Etling, & Brekke, 1996). Overconfidence constitutes a potential reason for the failure of forewarnings because it may affect people's awareness about their vulnerability to influences. As experts are particularly overconfident when making judgments (Törngren & Montgomery, 2004), they are expected to be even less responsive to forewarnings than non-experts in reducing the magnitude of the anchoring effect.

In Studies 1 to 3, expertise was operationalized through different experience levels. Students either had experience with an anchoring task (e.g., estimation of a flat rent) or were unfamiliar with it (e.g., estimation of a company value). In Study 4, experts (management consultants) and non-experts (students) were directly compared while performing economic anchoring tasks. First, half of the participants received an overconfidence note (Study 1) or their confidence was reduced by difficult general knowledge questions (Studies 2 to 4). Afterwards, half of the participants received a forewarning about the anchoring effect. As predicted, forewarnings reduced the magnitude of anchoring in the low expertise conditions. More importantly, the magnitude of anchoring was also reduced in the high expertise conditions, but only after a confidence reduction (Studies 2 to 4), whereas an explicit overconfidence note (Study 1) failed to produce the same effect (Study 1). For management consultants who received a forewarning, the effect of the confidence reduction on the magnitude of anchoring was mediated by the degree of confidence (Study 4).

Taken together, higher confidence of experts can explain why expertise impedes the effectiveness of forewarnings about anchoring effects. However, reducing experts' confidence can render a forewarning effective even in the case of high expertise. Directions for future research and practical implications are discussed.

DEUTSCHE KURZZUSAMMENFASSUNG

Ankereffekte sind robust und schwer zu korrigieren. Vorwarnungen über den Einfluss von Ankerwerten konnten bisher zu keiner Reduktion des Ankereffektes führen (Wilson et al., 1996). Ein möglicher Grund könnte die Überschätzung der eigenen Urteilsfähigkeit und damit die Unterschätzung der eigenen Betroffenheit von Urteilsheuristiken sein. Da die Überschätzung der eigenen Urteilsfähigkeiten bei Experten besonders ausgeprägt ist (Törngren & Montgomery, 2004), könnte es sein, dass Experten den Ankereffekt nach einer Vorwarnung weniger korrigieren als Nicht-Experten.

In den Studien 1 bis 3 wurde die Variation von Expertise angestrebt, indem Studierende Ankeraufgaben bearbeiten mussten, in denen sie viel (z.B. Schätzung einer Wohnungsmiete) oder wenig Erfahrung hatten (z.B. Schätzung des Wertes eines Unternehmens). In Studie 4 wurden Unternehmensberater (Experten) mit Studenten (Nicht-Experten) bezüglich Ankeraufgaben mit ökonomischen Inhalten verglichen. Am Anfang der Studien wurde die Hälfte der Probanden auf eine mögliche Überschätzung ihrer Urteilsfähigkeit hingewiesen (Studie 1) oder die Einschätzung ihrer Urteilsfähigkeit durch die Bearbeitung von schwierigen Allgemeinwissensaufgaben reduziert (Studien 2 bis 4). Danach erhielt die Hälfte der Probanden vor der Bearbeitung der Aufgaben eine Vorwarnung zum Ankereffekt (Studien 1 bis 4). Wie angenommen, konnten Vorwarnungen den Ankereffekt bei geringer Expertise reduzieren. Darüber hinaus war die Vorwarnung auch bei hoher Expertise erfolgreich, wenn vorher auch die Einschätzung der Urteilsfähigkeit experimentell reduziert worden war (Studien 2 bis 4). Ein Hinweis über das Risiko der Überschätzung der eigenen Urteilsfähigkeit hatte dagegen keinen Effekt (Studie 1). Für Unternehmensberater in der Bedingung mit Vorwarnung konnte gezeigt werden, dass die Wirkung der

experimentellen Verringerung der Einschätzung der Urteilsfähigkeit auf das Ausmaß des Ankereffektes von der Urteilssicherheit mediiert wurde (Studie 4).

Insgesamt kann die erhöhte Einschätzung der eigenen Urteilsfähigkeit bei hoher Expertise erklären, warum Expertise die Wirksamkeit einer Ankeraufklärung vermindert. Eine experimentelle Verringerung der Einschätzung der Urteilsfähigkeit von Experten verhindert diese negativen Auswirkungen von Expertise und macht Vorwarnungen auch bei hoher Expertise wirksam. Praktische Implikationen und sich ergebende Richtungen für zukünftige Forschung werden diskutiert.

THEORETICAL BACKGROUND

In 2007, the American housing market bubble burst, triggering the biggest financial and economic crisis since World War II. The seemingly sudden and rapid collapse is often compared to a natural disaster, implying the inevitability of this crisis. However, increasing evidence suggests that substantial concerns did previously exist regarding complicated financial products and the exaggerations of the prices in the U.S. housing market (Rajan, 2005; Shiller, 2006; Taleb, 2006). Indeed, analyses in the aftermath of the crisis revealed that expert decision makers were explicitly warned about the upcoming disaster. For example, economist Nouriel Roubini presented his prognoses at a conference of the International Monetary Fund in Washington in 2006: He predicted a breakdown of the U.S. real estate market, trillions of dollars in bad mortgage loans, and collapses of banks and hedge funds (Roubini, 2010). These forewarnings almost perfectly predicted the allegedly unforeseeable events.

In line with the ignorance of the markets, psychological research has demonstrated that decision makers often fail to benefit from forewarnings about biases, such as, the *halo effect* (Wetzel, Wilson, & Kort, 1981), the *overconfidence effect* (Armor, 1999), and the *anchoring effect* (Wilson et al., 1996). The latter, which denotes the assimilation of numerical judgments under uncertainty towards a given numerical standard, is a particularly robust bias (Chapman & Johnson, 1999; Mussweiler & Strack, 1999) and cannot be reduced by forewarnings (Wilson et al., 1996). At the same time, it occurs in many essential real-life settings such as negotiations (Galinsky & Mussweiler, 2001), product purchase (Mussweiler, Strack, & Pfeiffer, 2000), or in the courtroom (Englich, 2006; Englich, Mussweiler, & Strack, 2005). It therefore seems particularly important to investigate why anchoring remains largely unaffected by forewarnings and to find ways to improve the effectiveness of these forewarnings.

In a study by Wilson et al. (1996), participants were forewarned about the anchoring effect. These forewarnings were unable to reduce the anchoring effect, even though they were given just before the anchoring task and included an example similar to the task itself. Why were these forewarnings about anchoring so stunningly ineffective, just like the alarming signs before the financial crisis? The failure to correct the anchoring effect by a forewarning was explained by the participants' belief that their own judgments were less biased than those of others (Wilson et al., 1996, p. 398). The forewarning proved to be successful in convincing participants that people in general are susceptible to anchoring effects, but they were not willing to assume bias in their own judgments. In short, participants considered themselves immune to the anchoring effect.

As important decisions are often made by experts in their specific domain, it is highly relevant to consider the impact forewarnings may have on their decision making process. On the one hand, there are reasons to believe that experts may debias more effectively, for instance, they typically have access to more anchor-inconsistent information (c.f., Englich, 2008), which can reduce anchoring effects (Mussweiler et al., 2000). On the other hand, it has been shown that experts are often particularly overconfident (e.g., Englich, Mussweiler, & Strack, 2006; Törngren & Montgomery, 2004; Trafimow & Sniezek, 1994). If, as Wilson et al. (1996) suppose, overestimating the validity of one's own judgments is, in fact, the reason why forewarnings fail, expertise may affect debiasing more negatively than positively. For example, a professional investor, whose investing decisions are biased by past stock prices (Mussweiler & Schneller, 2003), could be overconfident about his decisions because of his high expertise level. He might therefore leave this bias uncorrected, despite being forewarned about the anchoring effect. As it can be seen in this example, the effects of expertise on debiasing attempts are of practical importance. Nevertheless, this issue has so far not been examined in the context of anchoring or other cognitive biases.

To elaborate my hypotheses in a more detailed manner, I will first provide a short overview of the robustness of the anchoring effect and the effectiveness of debiasing measures in different types of anchoring. This is followed by an introduction of the concept of overconfidence and a description of the risk that overconfidence may entail for the success of debiasing measures. Finally, I will discuss the particular role that overconfidence may occupy in high-expertise settings.

The robustness of anchoring

Numerical standards (anchors) can assimilate estimations of unknown quantities (Tversky & Kahneman, 1974). This phenomenon is called the *anchoring effect*. It is best illustrated by a prominent example: In the classical study by Tversky and Kahnemann (1974), participants estimated the percentage of African countries in the United Nations (UN), which was influenced by a number which was generated by a wheel of fortune. First, participants were asked a comparative question, that is, whether the percentage of African nations in the UN is higher or lower than an anchor (65% vs. 10%), which had obviously been picked by spinning a wheel of fortune. In a subsequent absolute anchoring question, participants gave their estimate of this percentage. These absolute judgments were significantly assimilated towards the random number; hence, the mean estimation of participants who had obtained the high anchor was 45%, as compared to 25% for participants who had obtained the low anchor.

Such demonstrations of the anchoring effect under the controlled conditions of a laboratory are abundant (e.g., Chapman & Johnson, 1999; Epley & Gilovich, 2001; Janiszewski & Uy, 2008; Mussweiler & Strack, 2000a). Additionally, the practical implications of the anchoring effect have also been demonstrated in numerous real-life settings (Chapman & Bornstein, 1996; Galinsky & Mussweiler, 2001; Kaustia, Alho, & Puttonen, 2008; LeBoeuf & Shafir, 2009; Mussweiler et al., 2000; Northcraft & Neale, 1987; Stewart, 2009). For example, in face-to-face and E-mail negotiations, the party that set an anchor by making the first offer received a better outcome (Galinsky & Mussweiler, 2001; Ritov, 1996). Moreover, first offers were strong predictors of counteroffers and final settlement prices (Galinsky & Mussweiler, 2001). However, the anchoring effect not only affects judgments of non-experts, but also those of experienced judicial experts, who let their sentencing decisions be guided by random numbers (Englich et al., 2005, 2006; Englich & Soder, 2009). The same is the case for real estate agents whose estimations of a property value were biased by manipulated listing prices (Northcraft & Neale, 1987).

Anchoring is also one of the most robust cognitive biases (Chapman & Johnson, 1999; Mussweiler & Strack, 1999). It occurs even if the anchor values are irrelevant for the absolute estimate, because they were obviously selected at random (e.g., Englich et al., 2006; Mussweiler, 2001; Mussweiler & Strack, 2000b; Tversky & Kahneman, 1974). Additionally, anchoring is not mitigated by extreme and implausible anchors (e.g., Mussweiler & Strack, 2001a; Strack & Mussweiler, 1997). For example, estimates of the annual mean temperature in the Antarctic were assimilated to both an unreasonably high anchor value of 700 °C and to a plausible high anchor value of -17 °C (Mussweiler &

Strack, 2001a). Moreover, anchoring effects persist over time. In a study conducted by Mussweiler (2001), anchoring effects still appeared one week after the anchor value had been considered. Probably the most striking evidence of the robustness of this phenomenon, however, stems from research using anchors which are presented incidentally (Critcher & Gilovich, 2008) or even outside the participant's awareness (Mussweiler & Englich, 2005; Reitsma-van Rooijen & Daamen, 2006). For example, high or low anchor values influenced participants asked to evaluate the cost of an average German car even when they were presented subliminally while participants worked out their estimations (Mussweiler & Englich, 2005). Even a coincidental number displayed on athletes' jerseys has proved to bias the judgments of their performance (Critcher & Gilovich, 2008). As we are almost constantly surrounded by random numbers, as in advertisements, media, or public signs, one might conclude that anchors affect most numerical estimations.

Debiasing of the anchoring effect: Distinctions, mechanisms, and methods

There is wide agreement across different psychological and non-psychological disciplines (economics, law, educational science) that human thinking and decision making are not as rational as once commonly postulated (Dawes, 2001; Kahnemann & Tversky, 1996; Simon 1955, 1957; Stanovich & West, 2000; Tversky & Kahneman, 1974). Many cognitive biases have been discovered – Krueger and Funder (2004) list 42 biases in their review – and there is a vast body of research on moderators and processes. In remarkable contrast to the abundant research on cognitive biases in human judgment, debiasing has been a relatively neglected empirical issue (cf., Arkes, 1991; Larrick, 2004; Petty & Wegener, 1993). Furthermore, researchers have made far more progress in cataloging cognitive biases than in discovering ways to correct or prevent them (Lilienfeld, Amirati, & Landfield, 2009). Accordingly, the terms "debias" or "debiasing" only yield 128 references in a PsychINFO search (June 24, 2011) – compared to 4,835 references for the terms "cognitive bias" or "cognitive biases". As discussed by Lilienfeld et al. (2009), psychology could make a valuable contribution to society by developing effective and applicable debiasing measures.

With regard to the considerable practical implications (Chapman & Bornstein, 1996; Galinsky & Mussweiler, 2001; Kaustia, Alho, & Puttonen, 2008; Mussweiler et al., 2000; Northcraft & Neale, 1987; Stewart, 2009) and the perseverance of the anchoring effect (Chapman & Johnson, 1999; Mussweiler & Strack, 1999), it seems particularly important to find effective methods of reducing it. However, presenting solutions to reduce this bias is complicated by the fact that anchoring effects seem to be multifaceted. Empirical evidence suggests that different types of anchoring effects exist. Even more importantly for my research, the effectiveness of different debiasing methods varies depending on the type of anchoring effect considered (Englich, 2008; Epley & Gilovich, 2001, 2004, 2005, 2006).

The first important distinction is based on the degree of attention directed to the anchor value. *Standard anchoring effects* involve paradigms in which the person devotes some thought to the anchor value. Typically, participants are asked to give a comparative judgment about the target quantity, followed by an absolute judgment. In the above-mentioned study by Tversky and Kahnemann (1974), they were first asked whether the percentage of African nations in the UN is higher or lower than a presented anchor (comparative judgment) and then gave their best estimates of the value (absolute judgment). *Basic anchoring effects*, on the other hand, involve paradigms with no direct comparison of anchor and target (e.g., Wilson et al., 1996, Study 3). For example, Wilson et al. (1996, Study 3) asked judges to copy either five pages of high numbers or five pages of irrelevant words before giving their judgments. In the subsequent estimation task, writing high numbers resulted in higher estimates. *Basic anchoring* is the most fragile type of anchoring (Brewer & Chapman, 2002) and can be reduced by knowledge (Englich, 2008). Therefore, in my attempt to reduce anchoring effects I focus on the more persistent *standard anchoring effects* in order to make a stronger point for my hypotheses.

A second important categorization differentiates between anchoring paradigms with e*xternally provided anchors* and *self-generated anchors*. In the *externally provided anchoring paradigm*, the anchors are explicitly (e.g., Tversky & Kahnemann, 1974) or implicitly (e.g., Wilson et al., 1996) provided by the experimenter, before participants make their absolute estimate. In the *self-generated anchors paradigm*, the comparative standards are generated by the participants themselves (Epley & Gilovich, 2001, 2004, 2005, 2006). For example, in this paradigm, participants are asked to estimate the freezing point of vodka. It is assumed that they start out with a self-generated anchor of 0°C as the freezing point of water and then adjust downwards knowing that the freezing point of alcohol is lower (e.g., Epley & Gilovich, 2006).

The anchoring effect can be reduced either by general debiasing strategies like forewarnings or incentives (Epley & Gilovich, 2001, 2005, 2006; Wilson et al.; 1996) or by interventions which use research about specific processes to debias (e.g., Mussweiler et al., 2000; Simmons, LeBoeuf, & Nelson, 2010). Additionally, different explanations for the anchoring effect predict different general debiasing strategies to be successful.

Therefore, it is important to be aware of the processes which explain the anchoring effect. To date, six explanations have been proposed to account for the anchoring effect: (1) conversational inferences, (2) numerical priming, (3) insufficient adjustment, (4) scale distortion, (5) selective accessibility, and (6) elaboration-based view of anchoring. The following sections elaborate further on these explanations.

Conversational inferences

According to the account of conversational inferences, participants who anticipate the experimenter to be maximally informative (Grice, 1975) may assume that the provided anchor value is close to the true value and consequently assimilate their estimation to it (Jacowitz & Kahnemann, 1995; Schwarz, 1994).

Numerical priming

A second account proposes that anchoring effects are purely numeric (Jacowitz & Kahneman, 1995; Wilson et al., 1996; Wong & Kwong, 2000). According to this approach, an anchoring task simply makes the anchor value more accessible, so that this value is likely to influence the subsequent absolute judgment (Mussweiler, Englich, & Strack, 2004).

Insufficient adjustment

The anchoring and adjustment hypothesis describes anchoring as a process by which people anchor on a given standard and then adjust their initial judgment until they reach the boundary of plausible values for the estimate (Tversky & Kahneman, 1974). The extent of adjustment may still be insufficient. For example, participants who are asked to estimate the freezing point of vodka, may use a value of 0 °C as a starting point. Then they may determine whether this value is too low or too high, and assimilate in the appropriate direction until the first plausible – but not necessarily correct – value is found.

Scale Distortion

The scale distortion theory explains anchoring by a shift in the use of the response scale (Frederick & Mochon, 2011). The underlying representation of the judgment target is assumed to remain stable. Accordingly, it should be of no relevance if the targets in the comparative versus the absolute question are identical or not. For instance, absolute estimations of the weight of an adult wolf are assimilated towards the anchor, irrespective of whether the comparative question contains a wolf or a sea turtle as comparative standard (Frederick & Mochon, 2011).

Selective accessibility

The selective accessibility model (SAM) of anchoring (Mussweiler & Strack, 1999, 2001b; Strack & Mussweiler, 1997; for a related account, see Chapman & Johnson, 1994, 1999) proposes that anchoring is essentially a knowledge accessibility effect. It involves two fundamental principles: (1) hypothesis-consistent testing and (2) semantic priming. Similar to the process of hypothesis-consistent testing, the model assumes that the comparison of the target to the anchor value in a comparative question changes the accessibility of knowledge about the target. More specifically, the accessibility of anchor-consistent knowledge is selectively increased. In line with this theory, asking whether more or less than 10% of the UN nations are African would lead participants to evaluate this option as if it were the true value. In doing so, they selectively activate knowledge that is consistent with this assumption and thereby allow the anchor to exert its influence. (e.g., “Asia is far bigger than Africa”).

Elaboration-based view of anchoring

The previously presented explanations are mostly not seen as mutually exclusive, but as complementary (Mussweiler et al., 2004; Mussweiler & Strack, 1999, 2000; Wilson et al., 1996).

An attempt to combine different accounts explaining anchoring in one coherent theory has been recently made (Wegener, Petty, Blankenship, & Detweiler-Bedell, 2010a)[1]. It was proposed that anchors can adopt multiple roles (Blankenship et al., 2008; Wegener, Petty, Detweiler-Bedell, & Jarvis, 2001). More precisely, the anchoring effect may result from relatively thoughtful processes (high elaboration) in some cases, and from relatively non-thoughtful processes (low elaboration) in others. According to persuasion theories (see Hovland, Janis, & Kelley, 1953), high elaboration is defined as an inspection of judgment-relevant information in light of existing knowledge and beliefs (Petty & Cacioppo, 1984). With respect to the domain of anchoring, particularly the processes proposed by the SAM (Mussweiler & Strack, 1999; Strack & Mussweiler, 1997) are seen as highly elaborative. In low elaboration settings, conversational inferences and numerical priming have been proposed as the mechanisms at work (Wegener et al., 2010).

[1] For a discussion of this account see Epley & Gilovich, 2010; Russo, 2010; Frederick, Kahnemann, & Mochon, 2010; Wegener, Petty, Blankenship, & Detweiler-Bedell, 2010b.

Specific debiasing strategies

As descibed above, some debiasing measures use these explanations to derive specific debiasing strategies. For example, the SAM predicts that anchoring is caused by the increased accessibility of *anchor-consistent* information. An activation of *anchor-inconsistent* information should therefore reduce the anchoring effect. Mussweiler et al. (2004) tested and confirmed this hypothesis. In the study, car mechanics were approached and asked for the price of a 10-year-old-car. Then they were informed about the estimate of the client himself – the anchor value ("I thought that the car should sell for about 2,800/5,000 Marks", Mussweiler et al., 2000, p. 1143). If participants were then asked to list arguments against this price, the anchoring effect was significantly reduced.

General debiasing strategies

The SAM and the insufficient adjustment account also make diverging predictions about the success of general debiasing strategies (e.g., forewarnings, incentives) (Epley & Gilovich, 2001, 2005, 2006; Mussweiler & Strack, 1999). While the extent of an insufficient adjustment should vary by the amount of effortful thinking, the increased accessibility of anchor-consistent information should be independent of it. In an *externally provided anchors* setting, the predictions of the SAM found broad empirical support (e.g., Chapman & Johnson, 2002; Epley & Gilovich, 2005; Tversky & Kahnemann, 1974; Wilson et al., 1996). Anchoring effects could not be reduced by forewarnings (Wilson et al., 1996; Epley & Gilovich, 2005), incentives (Chapman & Johnson, 2002; Epley &

Gilovich, 2005)[2] or time pressure (Mussweiler & Strack, 1999). In contrast, in a *self-generated anchors* setting forewarnings and incentives (e.g., financial incentives) reduced the effect of self-generated anchors (Epley & Gilovich, 2001, 2004, 2005, 2006).

The different results of forewarnings and incentives concerning externally provided or self-generated anchors are explained by the degree of automaticity in information processing (Epley & Gilovich, 2005, 2006). Self-generated anchors serve as starting points, but then undergo effortful and deliberate adjustment until the value seems right. In contrast, the process of knowledge activation through externally provided anchors tends to be largely automatic and is therefore unlikely to be affected by deliberate thought (Epley & Gilovich, 2005, 2006).

An alternative explanation for these different results concerning self-generated and externally provided anchors has recently been discussed (Simmons et al., 2010). Whereas people know the direction of the adjustment from a self-generated anchor, they are less certain about the direction of adjustment from an externally provided anchor. They often believe that their initial adjustment was too far from the externally provided anchors and reduce their adjustment if they are asked to reconsider their judgments (Simmons et al., 2010). In support of this reasoning, Simmons et al. (2010) showed that incentives also increase adjustment from externally provided anchors if participants are certain about the direction of adjustment (Studies 2, 3a, 3b, & 5).

---

[2] A different result was found in the study of Wright & Anderson (1989), where a marginally significant reduction due to incentives was found. Additionally, incentives reduce the anchoring effect if people are certain about the direction of adjustment (Simmons, LeBoeuf, & Nelson, 2010).

In this dissertation, I propose an additional, slightly different explanation, which focuses on the degree of awareness of the necessity of adjustment. An adjustment to the boundary of plausible values makes sense only if the anchor value falls outside the range of plausible values (Mussweiler et al., 2004). A value may be implausible for a participant because it is absurdly extreme, or because it is known to be incorrect. In a self-generated anchors setting, judges are aware that they have to adjust away from the anchor because they know – from the beginning – that the self-generated anchor is not the correct answer (Epley & Gilovich, 2001). For example, participants who self-generate the freezing point of water as an anchor in order to estimate the freezing point of vodka probably know that 0°C is an implausible value (Mussweiler et al., 2004). As a consequence, the anchoring and adjustment account is particularly applicable to explain anchoring in self-generated anchors paradigms. In contrast, externally provided anchors have to be considered the correct answer, even if only for a moment (Epley & Gilovich, 2001). Judges are hence not necessarily aware of the need to correct their judgments away from the randomly determined anchors.

Additionally, judges may also be less conscious of the necessity to correct externally provided anchors, because it can be aversive to accept a random external influence on one's own judgments. For judges who consider themselves to be rational decision makers, being biased by random numbers (e.g., Tversky & Kahnemann, 1974) or receiving advice from obviously incompetent counselors (e.g., Mussweiler et al, 2000) constitutes a self-esteem threat. Even if forewarned, people may therefore deny their own susceptibility to the anchoring effect because of their overconfident belief to be immune to bias. As an empirical support of this reasoning, participants expect others to be more prone to the anchoring effect than they themselves are (Wilson et al., 1996).

If these considerations are correct and overconfidence is in fact the reason why forewarnings are not successful in reducing the effect of externally provided anchors, forewarnings should be more effective if confidence is reduced.

Overconfidence as a risk for debiasing

Overconfidence (Fischhoff, Slovic & Lichtenstein, 1977) comprises three subtypes (Moore & Healy, 2008). The first and best-researched (Moore & Swift, 2011) subtype is *overestimation*. People subjectively perceive that their own performance is better than it really is. For example, marketing students overestimate their performance on academic exams (Clayson, 2005), or investors overestimate the quality of their own investment decisions (Moore, Kurtzberg, Fox, & Bazerman, 1999). The second type of overconfidence is *overprecision*. People are excessively certain about the accuracy of their estimations. In knowledge tasks, for example, they choose too narrow confidence intervals when asked to determine an upper and a lower bound of their answers (Klayman, Soll, Gonzalez-Vallejo, & Barlas, 1999; Soll & Klayman, 2004). The third type is called *overplacement* (Larrick, Burson, & Soll, 2007). People regard themselves as better than they really are if they compare themselves to others. For example, the majority of American and Swedish students think that they are far more skillful and less risky drivers than the average driver (Svenson, 1981). About 88% of the US student group and 77% of the Swedish group believed themselves to be safer drivers than the median, which is objectively statistically impossible.

The overplacement subtype of the overconfidence bias, in particular, appears to be related to people’s conviction that they are less vulnerable to the anchoring effect than others (c.f., Wilson et al., 1996). The gap between accepting that cognitive biases generally exist while denying that people themselves are affected by them has also been discovered for other biases and is called the *bias blind spot* (Kruger & Gilovich, 1999; Pronin, Gilovich, & Ross 2004; Pronin & Kugler, 2007). The *bias blind spot* is defined as the bias of seeing oneself as less susceptible to biases than others (Pronin & Kugler, 2007; Pronin, 2008). People accept the general fact that cognitive biases are skewing human judgments – but they underestimate the impact of biases on their own behavior.

Negative effects of overconfidence on debiasing measures have already been shown in different domains. For example, overconfidence in the quality of intuitive judgments contributes to the reluctance to use helpful actuarial judgment aids (Sieck & Arkes, 2005). Like decision aids, forewarnings are also an external support which can be used or neglected depending on the level of overconfidence. As overconfidence is a ubiquitous phenomenon (Plous, 1995; Moore & Swift, 2011) and there is no apparent reason to believe that these negative effects of overconfidence are unique to the domain of decision aids, it seems logical to assume that overconfidence may undermine the effectiveness of an anchoring forewarning.

If this reasoning is true, a reduction of overconfidence should make an anchoring forewarning more effective. A reduction of overconfidence has often been shown to be difficult (Armor, 1999; Sharp, Cutler, & Penrod, 1988; Sieck & Arkes, 2005, Studies 1 & 2) and the success of various techniques is often restricted to a specific context (cf., Fischhoff, 1982). Nevertheless, a successful reduction of overconfidence has been accomplished by calibration feedback (e.g., Lichtenstein & Fischhoff, 1980; Sieck & Arkes, 2005, Study 3), considering the opposite (Koriat, Lichtenstein, & Fischhoff, 1980), difficult tasks at the beginning of the study (Trafimow & Sniezek, 1994), anticipation of a group discussion (Arkes, Christensen, Lai, & Blumer, 1987, Study 2), or asking participants seemingly "easy" questions which were actually more difficult and giving them feedback about the correct responses (Arkes et al., 1987, Study 2). In the latter study, participants were asked general knowledge questions which appeared to be easy but contained hidden difficulties. For instance, participants were asked which country is larger, Greenland or Great Britain. For several reasons, such as, for example, Great Britain is more available for participants (availability bias) or more populated (halo effect), participants gave incorrect responses, hardly doubting the correctness of their responses. After receiving feedback about the correct solution, these participants were less overconfident than participants in the control condition without hidden difficulties (Arkes et al., 1987, Study 2).

Experts – particularly confident

The decisions which led to the financial crisis were made by experts in their fields. As described above, politicians and managers did not react to forewarnings (Rajan, 2005; Shiller, 2006; Taleb, 2006). Was this a singular event, or are experts in general less receptive to forewarnings? This question is all the more important, because a lot of far reaching decisions in our society are made by experts. Politicians and managers are responsible for decisions which can cause wars (recently in Afghanistan or Iraq) or maximum credible accidents (recently in Fukushima). In everyday life, managers, doctors, and judges constantly face countless important decisions.

It seems reasonable that expertise has detrimental effects on the different stages of the debiasing process (Wilson & Brekke, 1994), as depicted in Figure 1. On the one hand, it is plausible to assume that experts should be able to correct for judgmental biases like the anchoring effect more effectively than non-experts for several reasons. First, experts know more about the target and should therefore have more anchor-inconsistent information at their disposal (cf., Englich, 2008). As the consideration of anchor-inconsistent information can reduce anchoring effects (Blankenship, Petty, Detweiler-Bedell, & Macey, 2008; Mussweiler et al., 2000), experts might more often overcome its influence. Second, in light of the insufficient adjustment approach (Epley & Gilovich, 2001, 2004, 2005, 2006), experts might consider a narrower boundary of plausible responses and thus move further away from the anchor. Third, experts may also be more aware of the direction and magnitude of a potential bias due to their broader and more profound knowledge in their specific domain of expertise. According to Wilson and Brekke (1994) and Simmons et al. (2010), these are important pre-requisites for successful correction. As a consequence, experts might correct more easily.

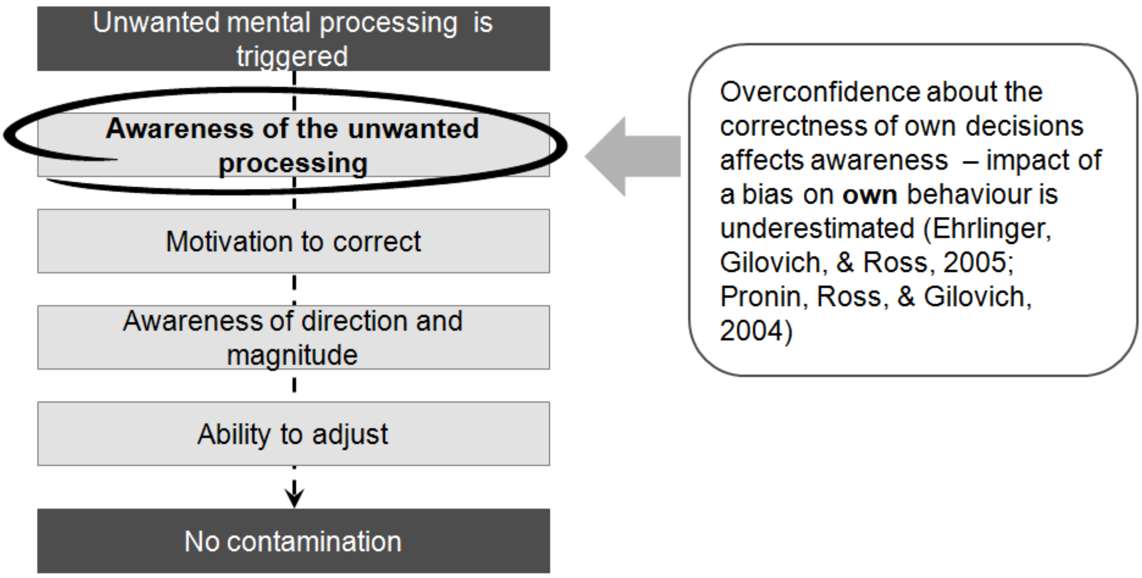

Figure 1: According to the presented reasoning, the process of mental contamination and mental correction (adapted from Wilson & Brekke, 1994) is disturbed by overconfidence in the stage of awareness about the unwanted processing.

On the other hand, psychological research suggests that experts are often as vulnerable to cognitive biases as non-experts (Englich & Mussweiler, 2001; Englich et al., 2005, 2006; Mussweiler et al., 2000; Northcraft & Neale, 1987; Whyte & Sebenius, 1997). Additionally, expertise has been shown to increase certainty about one's judgment, but leave bias undiminished (Englich et al., 2006; Törngren & Montgomery, 2004; Trafimow & Sniezek, 1994)[3]. For example, Northcraft and Neale (1987) let real estate agents and students estimate the value of a property. Although experts were equally biased by the anchoring effect in their professional domain as laymen, they were less aware of the biasing influences. In a similar vein, people have been shown to be more overconfident for tasks when they possess a self-declared expertise (Heath & Tversky, 1990), but their overconfidence decreases for tasks where they regard themselves as incompetent (Kruger, 1999). If somebody is convinced about the correctness of his decisions, she or he will probably be less aware of the fact that undesired and irrational influences interfere to a significant degree. It is therefore likely that experts are ignorant of their own susceptibility to biases, which is an important precondition for debiasing (Wilson & Brekke, 1994), see Figure 1.

In a related context, Arkes, Dawes, and Christensen (1986) explored the role of overconfidence and expertise in decision aid neglect. They first recorded participants' self-assessed knowledge about baseball rules, and then asked them to indicate which one out of three baseball players had won the Most Valuable Player (MVP) award. The researchers provided participants with the most important information for each player in

[3] For diverging results see Mckenzie, Liersch, & Yaniv, 2008. In a task that asked participants to estimate confidence intervals, experts estimated narrower intervals but the intervals were also nearer to the true values and therefore more often contained the true value. Thus, there was no net effect of hit rate and overconfidence.

the respective year. They also introduced them to a useful decision rule. Those participants who scored poorly on the quiz relied more on the decision rule during the MVP selection task, and so significantly outperformed the more experienced "baseball experts". The "expert" group, however, was more confident about their performance on the task. These results suggest that the "expert" group's overconfidence lowered their reliance on the aid, and thus impaired their decision quality. Although a decision aid is not entirely comparable to a forewarning, this result suggests that experts are more resistant than non-experts against embracing the potential benefits of external support in the form of a forewarning.

Taken together, experts typically feel more confident about their judgments than laymen do, while in reality they are often just as strongly affected by biases. By refusing to accept their own susceptibility to bias, they are likely to profit less from forewarnings.

The Current Research

The primary objective of the here presented studies is to determine the main pre-conditions which hinder or facilitate the correction of the anchoring effect. More precisely, they are designed to show that these pre-conditions differ systematically for experts and non-experts. As outlined above, experts tend to be particularly overconfident about their decisions and estimates (e.g., Englich et al., 2006; Törngren & Montgomery, 2004; Trafimow & Sniezek, 1994). As a consequence, experts should be less responsive to a forewarning about the anchoring effect compared to non-experts and therefore correct the anchoring effect less after this forewarning than non-experts. In order to show that overconfidence is indeed the reason of experts' unsuccessful bias correction after a forewarning, different confidence reduction methods are utilized in some of the experimental conditions. If the here presented reasoning is right, these confidence reductions should enable experts to correct for the anchoring effect after a forewarning.

Taken together, drawing on the example of experts' decision making, the following studies are designed to demonstrate that heightened confidence is a crucial factor that may hinder decision makers to eliminate anchoring effects after being forewarned.

EMPIRICAL EVIDENCE

Study 1

In Study 1, based on a student sample, experience was used as a proxy for expertise (c.f., Arkes et al., 1986; Hinds, 1999; for an overview, see Chi, 2006). Two estimation tasks were implemented as the main dependent variables: One task was closely related to students' everyday life, while participants were far less experienced with the other task.

Because of increased confidence in the high experience task, a forewarning about the anchoring effect was expected not to be effective, whereas it should be effective in the less familiar task. As a test of whether confidence is indeed the critical variable which moderates the effectiveness of an anchoring forewarning, confidence levels were varied. This was achieved by informing some participants about people's tendency to be overconfident about their own performance.

To test the above-mentioned hypotheses, a 2 (expertise: high vs. low) x 2 (overconfidence note: overconfidence note vs. no overconfidence note) x 2 (forewarning: forewarning vs. no forewarning) x 2 (anchor: low vs. high) mixed design with expertise as a within-factor and confidence reduction, forewarning, and anchor as between-factors was applied.

Method

Procedure

After being instructed to turn off their cell-phones and to follow the instructions carefully, participants were asked to complete a computer-based questionnaire in a laboratory on the university campus. The questionnaire was designed using Unipark EFS Survey. Its four parts were administered in the following order: The note about people's tendency to be overconfident (for half of the participants), the anchoring forewarning (for half of the participants), two estimation tasks containing externally provided numerical standards (anchors) in a standard anchoring paradigm, and a few additional self-report questions. Participants were randomly assigned to one of the eight conditions resulting from the three between-factors, namely overconfidence note, forewarning, and anchor. Additionally, the order of the two estimation tasks (high vs. low expertise task) was counterbalanced across participants to avoid order effects. After finishing, all participants were thanked, debriefed, and given a chocolate bar as compensation.

Materials

Expertise. As experience served as a proxy for expertise (c.f., Arkes et al., 1986; Hinds, 1999; for an overview see Chi, 2006), each participant worked on one task involving a familiar topic and another task in which low levels of experience could be assumed. In the following, these tasks are denoted as high and low expertise tasks, because expertise is the concept which should be approximated by experience. In one task, the student participants were asked to estimate the rent of a specific shared flat in Cologne. The most important characteristics determining the monthly rent of this shared flat in Cologne (e.g., living space, location) were provided. Due to the fact that most students in Cologne live in shared flats, and the monthly flat rent comprises the major part of their income, they should be experienced in estimating adequate rental rates. The low expertise task consisted of estimating the value of a company. Again, the most important characteristics (e.g., turnover, industry cluster, etc.) needed to properly estimate the company's value were given. As student participants are rarely engaged in this type of estimation, their expertise for this task should be low.

Overconfidence note. Half of the participants received the following note which aimed at reducing confidence concerning their own decisions: "Numerous studies have shown that judgments and estimates are systematically biased by undesired influences. Even among well-educated and intelligent people, these influences can lead to deviations from purely rational and objective judgments and estimations. Typically, judges themselves do not notice these influences. This is because the extent of undesired influences is systematically underestimated and one's own capability to make objective decisions is systematically overestimated. Even if people know about such a bias, they often think that only other people are biased, but neglect their own vulnerability to the bias. This results in overrating the accuracy of one's own judgments and estimates."

Anchoring forewarning. Half of the participants received an explicit anchoring forewarning. Similar to the study by Wilson et al. (1996), the anchoring forewarning comprised an explanation and an illustrative example of the anchoring affect: "A well-known bias which frequently affects judgments and estimations is the anchoring effect. Anchoring means that a predetermined random number can serve as an anchor and can influence estimations of unknown quantities. The anchoring effect has been demonstrated in a broad variety of judgmental domains and under various conditions. For example, in a famous study on anchoring effects, the estimate of the percentage of African countries in the UN was influenced by an obviously random number generated by spinning a wheel of fortune. A high value on the wheel of fortune resulted in a higher estimation of the number of African states in the UN, whereas a lower value on the Wheel of Fortune led to a lower estimation."

Anchor. After providing them with the most important facts about the shared flat in Cologne (high expertise) and the company value (low expertise) respectively, participants were asked two subsequent questions, corresponding to the standard anchoring paradigm: First, they were asked to estimate whether the real value was lower, just right or higher than a randomly generated anchor ("Is this number too low, exactly correct, or too high?"). Second, they were requested to provide their absolute estimate on the given task. All participants completed both estimation tasks. The high and low anchor values were chosen on the basis of a pretest in which a comparable group of student participants (n = 59) gave estimates for the target quantities. High anchors were set at the 85th percentile of pre-test estimates, and correspondingly, low anchors were set at the 15th percentile. Resulting from the pretest, in the high expertise estimation task (shared flat rent), half of the participants received a high anchor (420 €), whereas the other half received a low anchor (220 €). Similarly, in the low expertise estimation task (company value estimation), half of the participants received a high anchor (22,000 million €), the other half received a low anchor (400 million €). The order of the tasks was randomized.

Additional measures. To verify the effectiveness of the expertise manipulation, participants indicated their subjective expertise in the estimates after each estimation task on the following dimensions: The lowest and highest conceivable values for an adequate estimate, subjectively perceived judgment certainty, thoroughness of the estimation, and personal judgment quality compared to other participants' estimates. Moreover, at the end of the experiment, participants answered a short questionnaire intended to determine whether the overconfidence note changed attention to the anchoring forewarning. First, participants answered five questions about the anchoring instruction and the awareness about the anchoring effect: familiarity with the anchoring effect before the experiment, comprehensibility of the forewarning, personal influenceability by the anchoring effect, attitude towards being affected by anchoring effects, and consideration of the anchoring forewarning while making the estimates. Finally, participants rated their mood and alertness, and answered questions about specific incidents during the experiment as well as their demographic data.

Participants

83 participants (47 female, M_{Age} = 23.5 years, SD_{Age} = 3.15 years) were recruited by approaching them on the university campus and asking them to take part in a study containing estimation tasks. To ensure a low expertise level for the low expertise task (estimation of a company value), students of economics and financial mathematics were excluded from participation. Participants were randomly assigned to one of the experimental conditions. Due to excessive acoustic disturbance or deficient understanding of the instructions, four participants (5 %) were excluded from further analyses.

Results

Manipulation check

Expertise. As previously supposed, the participants felt more confident in the high expertise task. The difference between the lowest and highest value which participants considered a possible estimate was significantly wider in the low expertise estimation task[4] (*M* = 1.52, *SD* = 2.61) than in the low expertise task (*M* = 0.38, *SD* = 0.22), $t(1, 78) = 3.78$, $p < .001$, reflecting less certainty about the correct value. Furthermore, participants indicated (all scales ranging from 1 to 9) that they felt more certain in the high expertise condition (*M* = 6.00, *SD* = 1.32) than in the high expertise condition (*M* = 3.09, *SD* = 1.70), $t(1, 78) = 12.32$, $p < .001$), rated their estimations as more thorough [*Ms* = 6.11 and 4.86, *SDs* = 1.43 and 1.38, $t(1, 78) = 7.07$, $p < .001$], and estimated their performance superior in comparison to the average peer [*Ms* = 56.34 % and 37.67 % in percentages, *SDs* = 17.05 and 18.35, $t(1, 78) = 7.12$, $p < .001$].

Overconfidence note. The overconfidence note did not have any significant effects on the self-report measures at the end of the study (consideration of the anchoring forewarning while forming one's own estimates, influenceability by the anchoring effect, attitude towards being affected by anchoring effects, mood, alertness), *ts* < 1.2.

Magnitude of the anchoring effect

In line with common practice, the standardized absolute estimates of the flat rent and the company value were analyzed (e.g., Epley & Gilovich, 2005, 2006; Strack &

[4] The difference between the lowest and highest value considered a possible estimate for each participant were divided by their estimates, to correct for the higher absolute values in the low expertise task.

Mussweiler, 1997) and corrected for outliers (van Selst & Jolicoeur, 1993)[5]. In the high anchor condition, higher z-values stand for stronger anchoring effects, whereas in the low anchor condition, higher z-values stand for weaker anchoring effects. For this reason, the z-values in the low anchor estimation tasks were reversed[6]. As a consequence, it was possible to aggregate the values across the high and low anchor conditions. To test my hypotheses, the standardized values of both estimations were submitted as repeated measures to a 2 (expertise: high vs. low) x 2 (overconfidence note: overconfidence note vs. no overconfidence note) x 2 (forewarning: forewarning vs. no forewarning) ANOVA, with expertise as a within-factor and overconfidence note and forewarning as between-factors.

In the low expertise task, the anchoring forewarning was expected to reduce the anchoring effect irrespective of the overconfidence note. In contrast, the anchoring forewarning was supposed to reduce the anchoring effect in the high expertise task only if confidence had been reduced by the overconfidence note prior to the forewarning.

The anchoring effect was significantly higher than zero across conditions (M = 0.53), $F(1, 75) = 68.02$, $p < .001$, $\eta^2_p = .48$. As predicted, the expertise x anchoring

---

[5] To minimize the disproportionate impact of outliers, a modified recursive outlier correction was used. A modified recursive outlier correction is less sensitive to sample size and amount of skewness than other outlier correction methods (van Selst & Jolicoeur, 1993). In a modified recursive outlier correction, the exclusion of extreme values below or above a certain cutoff criterion is repeated several times, until no extreme values remain or until the sample size of the restricted sample (including the temporarily excluded value) falls below four. As recommended by van Selst and Jolicoeur (1993), a cutoff criterion of 3.5 standard deviations was set. Outliers were replaced with group means.

[6] Due to this transformation, higher z-scores always indicate stronger anchoring effects in the here presented studies. Without the anchoring effect, the average z-score would be zero, because there would be no overall difference between high and low anchor. For example, if an estimation on a low anchor item was located between the low anchor and the overall mean (anchoring effect), it was transformed from a negative to a positive z-value by a multiplication with (-1).

forewarning two-way interaction was significant, $F(1, 75) = 9.50$, $p < .01$, $\eta^2_p = .11$. In the high expertise task, single contrasts showed no difference between the forewarning condition and the condition without forewarning (see Figure 2), $F(1, 75) < 2.7$, $p > .11$. In the low expertise task, however, the anchoring effect was significantly reduced by the anchoring forewarning (see Figure 2), $F(1, 75) = 7.20$, $p < .01$, $\eta^2_p = .09$. Deviating from my expectations, the interaction of expertise x overconfidence note x anchoring forewarning was not significant, $F < 1$. None of the remaining effects reached significance, all *Fs* < 1.7.

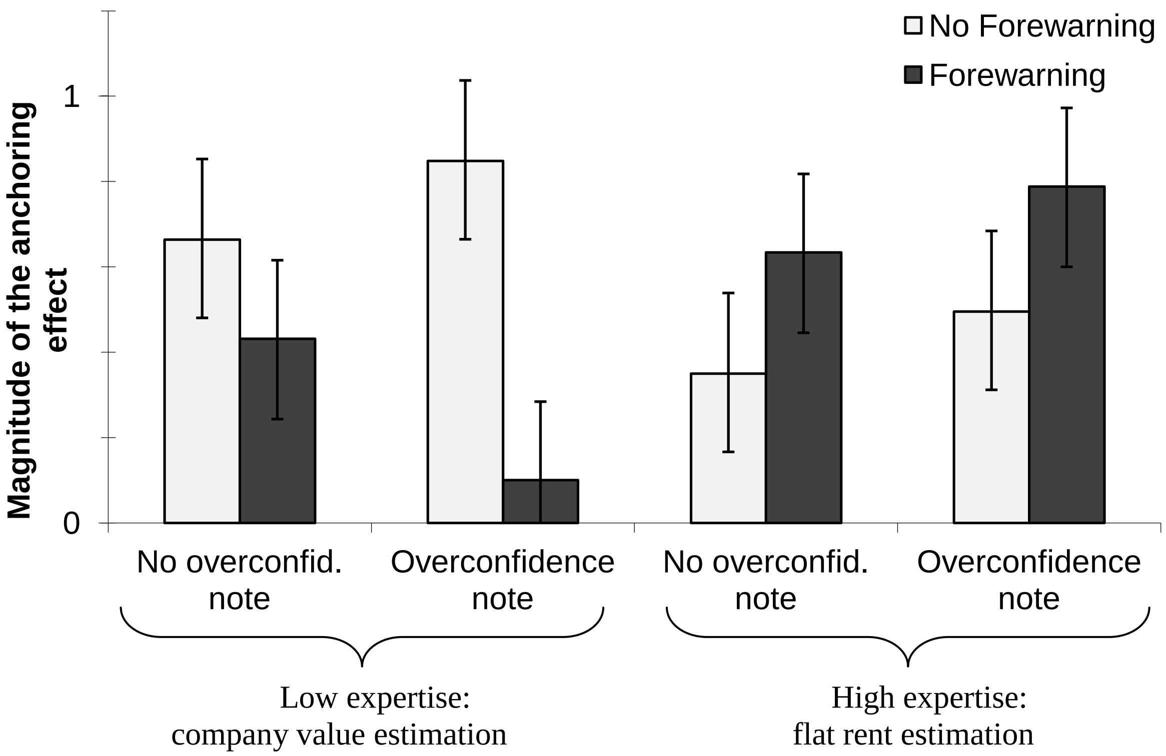

Figure 2: Magnitude of the anchoring effect (z-values with reverse coded low anchor items) by expertise (low vs. high), overconfidence note (no note vs. note), and anchoring forewarning (no forewarning vs. forewarning). Error bars represent standard errors. (Study 1).

Discussion

Taken together, the results show that a forewarning on anchoring effects can be an effective debiasing strategy if expertise is low. Differing from my expectations, the overconfidence note did not succeed in making the anchoring forewarning effective when expertise was high. In a similar way, all self-report measures concerning the perception and consideration of the anchoring forewarning were not influenced by the overconfidence note.

Possibly, the overconfidence note itself was not considered sufficiently by the participants because of their overconfidence, causing them to underestimate their own vulnerability to biases (Pronin & Kugler, 2007; Pronin, Gilovich, & Ross, 2004; Pronin, Lin, & Ross, 2002). If this reasoning is right, confidence cannot be effectively reduced by explaining overconfidence on an explicit level. In contrast, an effective confidence reduction should work on an implicit level.

Study 2

In the second study, overconfidence should be reduced by ten difficult general knowledge questions. The completion of difficult questions has been identified as a means to reduce confidence (Trafimow & Sniezek, 1994). Again, experience was varied as a proxy for expertise by the same two estimation tasks as in Study 1. As in Study 1, a 2 (expertise: high vs. low) x 2 (confidence reduction: reduction vs. no reduction) x 2 (forewarning: forewarning vs. no forewarning) x 2 (anchor: low vs. high) mixed design with expertise as a within-factor and confidence reduction, forewarning and anchor as between-factors was applied.

Method

Procedure

Participants completed an online questionnaire, which was designed with Unipark EFS Survey. The procedure was similar to Study 1. Again, participants were randomly assigned to one of the eight conditions, and the order of the estimation tasks was counterbalanced across participants.

Materials

Expertise. Participants worked on the same two estimation tasks as in Study 1.

Confidence reduction. Half of the participants received a treatment which was aimed at reducing confidence in their own decisions. The confidence reduction consisted of ten difficult general knowledge questions (e.g., "How long is the Nile River?"; "What is the diameter of the moon?"). For these questions, participants had to estimate confidence intervals. A reduction of confidence by difficult general knowledge questions has already been shown to be successful (Trafimow & Sniezek, 1994). Nevertheless, a pretest was conducted to ensure the confidence-reducing impact of these tasks.

Pretest of the confidence reduction

In a pretest (n = 41), it was verified that confidence was reduced significantly by solving ten difficult general knowledge questions, whereas other potential confound variables remained unaffected. More concretely, pretest-participants were asked before or after answering the questions to rate (a) their general aptitude in making estimates, (b) the probability that their estimates may be inaccurate, and (c) their own capabilities to give exact estimates in comparison to others. These three ratings measured the three facets of overconfidence defined by Moore and Swift (2011). In a fourth question, participants were asked (d) how much they expected external factors to bias their estimates, which assessed one aspect of the bias blind spot (Pronin, Lin, & Ross, 2002). Participants' ratings on (b) and (d) were reverse coded, because more confident participants were expected to be less conscious about giving inaccurate estimations and being biased by external influences.

Pretest participants' ratings were combined to a single confidence index (Cronbach's $\alpha = .80$). This index was significantly reduced when confidence was assessed after answering the questions compared to assessment before the treatment [*Ms* = -0.30 and 0.29 with and without confidence reduction, respectively, $t(1, 39) = 3.84$, $p < .001$].

To control for other potential effects of the general knowledge task, the German version of the *Multidimensional Mood Questionnaire* (Steyer, Schwenkmezger, Notz, & Eid, 1997) was employed, which measures three bipolar scales (good – bad mood, alertness – tiredness, and calmness – restlessness). Additionally, subjective frustration was measured with a supplementary question. T-tests revealed no significant impact of the ten general knowledge tasks on any of the scales or their subitems, nor on my supplementary question, all $ts < 1.3$. Hence, the ten difficult estimation tasks proved to be suitable for confidence reduction, whereas they did not affect participants' mood, alertness, restlessness, or frustration.

Anchoring forewarning and anchor. The anchoring forewarning and the anchor values were identical to those used in Study 1.

Additional measures. After each of the two estimation tasks, participants answered the same five questions about their subjective feeling of expertise as in Study 1. At the end, participants again answered a short questionnaire. Unlike in Study 1, participants first rated how much their own estimates might have been influenced by the two anchors, to measure how the confidence reduction changed the awareness about the anchoring effect. They then answered the same questions as in Study 1: the five questions about the perception of the anchoring forewarning and the awareness about the anchoring effect, one question about both their mood and alertness, general questions about specific incidents during the experiment and their demographic information.

Participants

115 participants (57 female, M_{Age} = 26.3 years, SD M_{Age} = 3.91 years) were recruited by approaching them on the university campus and asking them to take part in a study containing estimation tasks. To ensure a low expertise level for the low expertise task (estimation of a company value), students of economics and financial mathematics were excluded from participation. Participants were randomly assigned to one of the experimental conditions. Due to excessive acoustic disturbance, implausibly long completion times, repeated participation, or clearly deficient understanding of the instructions, five participants (4 %) were excluded from further analyses.

Results

Manipulation check

Expertise. Data confirmed the assumption that participants would feel more confident about their estimations in the high expertise task. The difference between the lowest and highest value which participants considered a possible estimate, was significantly wider in the low expertise estimation task[7] (M = 1.02, SD = 0.52) than in the high expertise task (M = 0.39, SD = 0.18), $t(1, 109) = 13.63$, $p < .001$]. Furthermore, participants indicated (all scales ranging from 1 to 9) that they felt more certain in the high expertise condition [*Ms* = 6.07 and 3.02, *SDs* = 1.43 and 1.70, for high and low expertise respectively, $t(1, 109) = 15.50$, $p < .001$], rated their estimations as more thorough [*Ms* = 6.09 and 4.91, *SDs* = 1.37 and 1.39, $t(1, 109) = 8.80$, $p < .001$], felt more competent [*Ms* = 5.72 and 2.85, *SDs* = 1.52 and 1.54, $t(1, 109) = 15.86$, $p < .001$], and regarded their performance as superior in comparison to the average peer [*Ms* = 54.99 % and 44.08 % in percentages, *SDs* = 19.60 and 21.38, $t(1, 109) = 4.49$, $p < .001$].

Confidence reduction. As expected, a confidence reduction should change how the anchoring forewarning is made use of, particularly in the high expertise task. More precisely, in the high expertise estimation task, participants were expected to become more aware that their own estimates might have been influenced by the anchors. In contrast, awareness about the influence of the anchors was not expected to increase significantly in

---

[7] As in Study 1, the difference between the lowest and highest value considered a possible estimate for each participant were divided by their estimates, to correct for the higher absolute values in the low expertise task.

the low expertise task, because it should be elevated in the first place as a result of low task confidence. In fact, a repeated-measures ANOVA of bias awareness with confidence reduction as between-factor and expertise as within-factor revealed that in the low expertise task, participants were generally more conscious of the influence the anchor exerted (M = 5.95, SD = 2.38), as opposed to the high expertise task (M = 4.95, SD = 2.20), $F(1, 108) = 12.38$, $p < .001$]. More importantly, a significant interaction between task expertise and confidence reduction occurred [$F(1, 108) = 5.44$, $p = .02$]. As expected, single contrasts show that the confidence reduction failed to moderate awareness of the influence of the anchor in the low expertise task [Contrast confidence reduction (reduction vs. no reduction): $F < 1$]. However, the confidence reduction significantly increased awareness about the influence of the anchor in the high expertise task [Awareness with confidence reduction (M = 4.48, SD = 2.16) vs. awareness without confidence reduction (M = 5.44, SD = 2.15), $F(1, 108) = 5.45$, $p = .02$].

Additionally, participants' self-reports indicated that they considered the anchoring forewarning significantly more while forming their own estimates after the confidence reduction [Ms = 5.46 and 4.00, SDs = 1.94 and 2.01 with and without confidence reduction, $t(1, 55) = 2.71$, $p = .01$]. The confidence reduction did not have any significant effects on other self-report measures (influenceability by the anchoring effect, attitude towards being influenced by anchoring effects, mood, alertness), $ts < 1.2$.

Magnitude of the anchoring effect

The procedure of outlier exclusion, z-standardization, and reverse-scoring of low anchor tasks was the same as in Study 1. The estimations for each expertise level were submitted as repeated measures to a 2 (expertise: high vs. low) x 2 (confidence reduction: reduction vs. no reduction) x 2 (forewarning: forewarning vs. no forewarning) ANOVA with expertise as a within-factor and confidence reduction and forewarning as between-factors. As expected, the anchoring effect was significantly higher than zero across conditions (M = 0.54), $F(1, 105) = 30.34$, $p < .001$, $\eta^2_p = .22$. Additionally, a marginally significant main effect of the anchoring forewarning was revealed. Anchoring effects on all expertise levels were marginally stronger without a forewarning (M = 0.69) than when a forewarning was provided (M = 0.40), $F(1, 105) = 3.78$, $p = .05$, see Figure 3. More interestingly, the expected expertise x confidence reduction x anchoring forewarning three-way interaction was marginally significant, $F(1, 105) = 3.5$, $p = .06$, $\eta^2_p = .03$. None of the remaining effects obtained significance, all *Fs* < 1.3. To better understand this interaction, the two estimations (estimation of a company value and a flat rent) were separately submitted to a 2 (confidence reduction: reduction vs. no reduction) x 2 (forewarning: forewarning vs. no forewarning) ANOVA.

Magnitude of the anchoring effect: company evaluation

In the low expertise estimation task, the anchoring forewarning was predicted to reduce the anchoring effect irrespective of the confidence reduction. In accord with the hypothesis, a preceding anchoring forewarning led to an anchoring effect that was marginally weaker (M = 0.40) than without the forewarning (M = 0.69), $F(1, 106) = 3.39$, $p = .07$, $\eta^2_p = .03$, see Figure 3.

Magnitude of the anchoring effect: flat rent

In the high expertise estimation task, the anchoring forewarning should reduce the anchoring effect only if confidence is reduced before the forewarning. Accordingly, the expected two-way interaction between confidence reduction and anchoring forewarning occurred, $F(1, 106) = 4.30$, $p = .04$, $\eta^2_p = .04$, see Figure 3. Single contrasts show that the anchoring forewarning failed to moderate the anchoring effect without a preceding confidence reduction in the high expertise task, $F < 1$. However, the anchoring effect is significantly moderated by an anchoring forewarning after a confidence reduction in the high expertise task, $F(1, 106) = 4.50$, $p = .04$, $\eta^2_p = .04$.

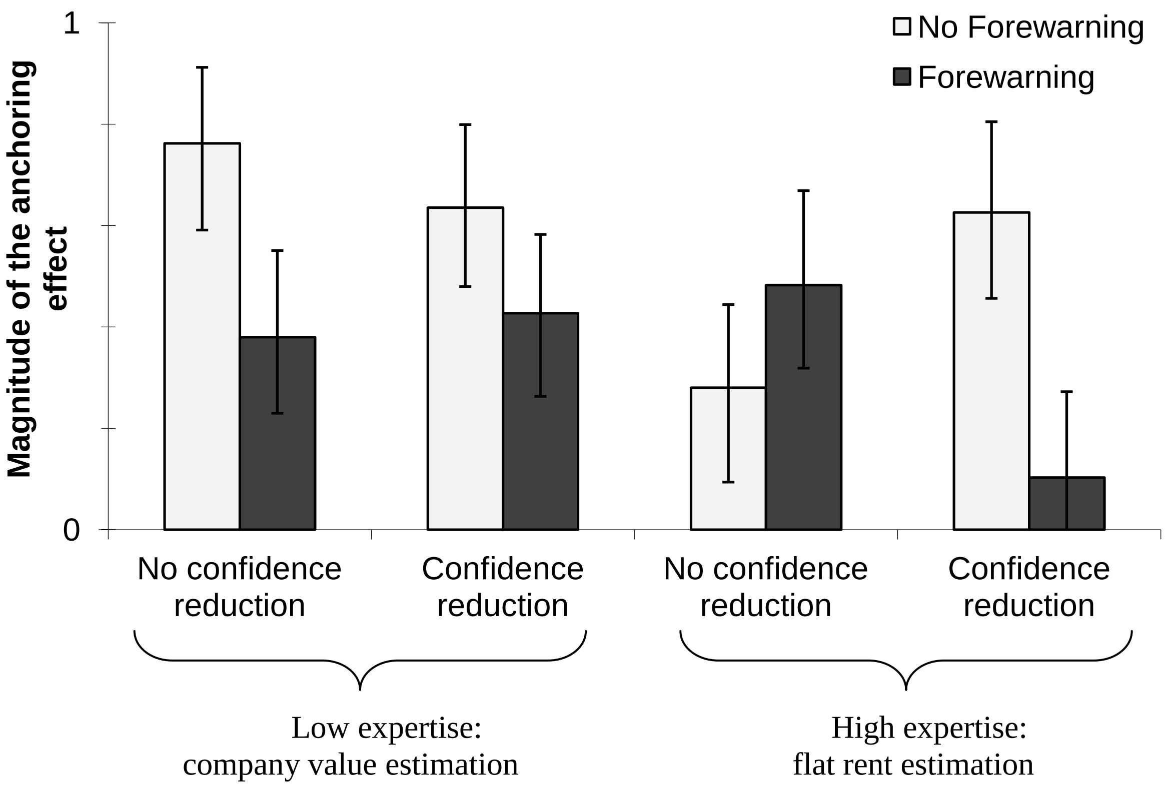

Figure 3: Magnitude of the anchoring effect (combined z-values with reverse coded low anchor items) by expertise (low vs. high), confidence reduction (no reduction vs. reduction), and forewarning (no forewarning vs. forewarning). Error bars represent standard errors. (Study 2).

In the high expertise task, the combination of the confidence reduction and the forewarning led to an anchoring effect that did not differ significantly from zero, $t < 1$, whereas in all other conditions, the anchoring effect was significantly higher than zero, all $ts > 8.55$, all $ps < .001$.

Discussion

First, Study 2 shows that a forewarning on anchoring effects can be an effective debiasing strategy if expertise is low. Moreover, if expertise is high, a forewarning can also reduce and even undo anchoring, given that confidence has initially been lowered. The effectiveness of a forewarning following the confidence reduction in the high expertise task indicates that excessive confidence may in fact be the reason why participants do not correct. As opposed to the explicit confidence reduction attempt in Study 1, the implicit method applied in Study 2 proved to be successful.

Self-report measures support this reasoning. As expected, confidence measures after the estimation tasks were significantly higher in the high expertise task. Additionally, participants experience a stronger influence of the anchor value and consider the anchoring forewarning to a further degree if confidence is initially reduced in a high expertise setting. In a low expertise setting, these effects did not appear. This empirical evidence supports my hypothesis that lacking awareness about one's own susceptibility to biases constitutes a reason for the failure of forewarnings in high expertise decision making.

Although the pretest showed that the confidence reduction lowered confidence levels as intended, while frustration, mood, as well as alertness were not influenced by the manipulation, a risk of another confounded variable still remains, namely fatigue.

Whereas participants in the confidence reduction condition had to work on ten difficult general knowledge questions, participants in the control group continued without a comparable task. Furthermore, my findings are only based on two specific estimation tasks. In the next study, I therefore used a different confidence reduction with a similarly effortful control treatment, as well as different estimation tasks.

Study 3

Having obtained first support for the hypothesis that heightened confidence is indeed a factor which hinders expert decision makers to eliminate anchoring effects after a forewarning, I aimed at replicating these findings of Study 2 with other confidence reduction methods and a different set of estimation tasks.

As in Study 2, I applied a 2 (expertise: high vs. low) x 2 (confidence reduction: reduction vs. no reduction) x 2 (forewarning: forewarning vs. no forewarning) x 2 (anchor: low vs. high) mixed design with expertise as a within-factor and confidence reduction, forewarning and anchor as between-factors.

Method

Procedure

Participants completed an online questionnaire, which was designed with Unipark EFS Survey. The procedure was similar to Study 2. Again, participants were randomly assigned to one of the eight conditions, and the order of the estimation tasks was counterbalanced across participants.

Materials

Expertise. Again**,** expertise was varied, but as opposed to Studies 1 and 2, each participant worked not only on one, but on two tasks for which she or he had a lot of experience and on two tasks in far less familiar domains.

One of the high expertise tasks was the same as in Study 2 and included the estimation of a shared flat rent in Cologne. In the second high expertise task, participants estimated the average lunch expenses of a student in the local university canteen. The most important characteristics determining the money spent for lunch (e.g., types of food, additional drinks) were described. As most of the students in Cologne eat in the university canteen quite regularly, they are experienced with this estimation. Likewise, one of the low expertise tasks was similar to that used in Study 2, requiring the estimation of a company value. In the second low expertise task, participants estimated the number of gas stations in Germany, which was designed as a task they had not been confronted with before. The task also provided a description of the most important characteristics determining the number of gas stations (e.g., population in Germany, number of cars in Germany).

Confidence reduction. Different to Study 2, confidence was reduced by letting participants experience their own fallibility. Similar to Arkes et al. (1987), participants were asked seemingly “easy” general knowledge questions and received feedback about the correct responses. The questions appeared to be easy but contained hidden difficulties. For instance, participants were asked whether the carnival event "Mardi Gras" is celebrated in Rio de Janeiro or in New Orleans. As carnival is often associated with Rio de Janeiro, participants feel confident that Rio is the right answer, even though New

Orleans is the actual solution. After receiving feedback about the correct responses, participants were expected to feel that they overestimated their own ability to identify the correct answer (Arkes et al., 1987).

In the control condition, the questions appeared to be and were in fact difficult (Arkes at al., 1987). One question, for instance, read as follows: What is the highest volcano on earth, Ojos del Salado or Guallatiri? Like in the experimental condition, participants were informed about the actually correct responses after completing the five tasks. However, these participants should not realize that they overestimated their knowledge because they already knew about the difficulties of the responses when they answered.

Pretest of the confidence reduction

The general knowledge questions of Arkes et al. (1987) were designed for an American sample and both the perceived and actual difficulty of most questions would differ for German participants. I therefore conducted a pretest with twenty new questions to identify five seemingly easy questions with hidden difficulties for the experimental condition, and five questions matching in apparent and actual complexity for the control condition. In the questions which were finally used in Study 3, participants rated the five questions in the experimental condition as significantly easier than those used in the control condition [Ms = 4.5 and 7.3 respectively, on a nine point scale, $t(1, 164) = 16.7$, $p < .001$]. As intended, the difficulty (number of correct responses) of the questions was constant in both groups [Ms = 2.7 and 2.8 with and without hidden difficulties, $t(1, 164) = 0.9$, $p = .38$].

Anchoring forewarning. Half of the participants received the same anchoring forewarning as in Studies 1 and 2.

Anchor. As in Studies 1 and 2, participants first answered a comparative question and then provided an absolute estimate. One of the high expertise tasks always included a high anchor (420 € for the shared flat rent and 6 € for the money spent for lunch), the other one a low anchor (220 € for the shared flat rent and 3 € for the money spent for lunch). Similarly, one of the low expertise tasks always included a high anchor (100 billion € for the estimation of a company value and 550,000 for the number of gas stations), while the other one included a low anchor (4 billion € for the estimation of a company value and 5000 for the number of gas stations). The order of the tasks was randomized. The high and low anchor values were set at the 85th percentile and at the 15th percentile on the basis of different pretests (n = 57 for the money spent for lunch, n = 59 for the flat rent, n = 65 for the company value, and n = 34 for the number of gas stations).

Additional measures. Due to high correlations of the four items measuring the subjective feeling of expertise in Studies 1 and 2, only one question assessed the experienced judgment certainty after each estimation task. The same five questions as in Study 1 and 2 where administered to measure whether the confidence reduction changed the perception of the anchoring forewarning and the awareness about the anchoring effect. Because of the increased number of tasks compared with Study 2, it would have been difficult for participants to rate at the end of the study how much their own estimates had been influenced in each task. For this reason these questions were excluded in Study 3.

Additionally, three questions about each participant's general confidence level were introduced, which were asked in blocks after the second and the fourth estimation task:

Participants general aptitude in estimation tasks, their performance in such estimation tasks compared to other participants, and the probability of a hidden task difficulty. Again, participants answered one question concerning their mood and one question about their frustration, general questions about specific incidents during the experiment, and gave their demographic data at the end of the study.

Participants

176 participants (114 female, M_{Age} = 27.1 years, SD_{Age} = 6.54 years) recruited from a pool for online data-collection were given the chance to win a gift voucher by enrolling in a lottery. Participants were randomly assigned to one of the experimental conditions. As in the preceding studies, students of economics and financial mathematics were excluded from participation. Due to implausibly long completion times, repeated participations, and clearly deficient understanding of the instructions, ten participants (6 %) were excluded from further analyses.

Results

Manipulation check

Expertise. Participants felt more certain after finishing a high expertise task (M = 5.95, SD = 1.13) compared to a low expertise task (M = 3.27, SD = 1.50) on a nine point scale, $t(1, 164) = 23.17$, $p < .001$.

Confidence reduction. Again, the confidence reduction was expected to change participants' attention to the anchoring forewarning because of their increased awareness about their own susceptibility. Correspondingly, they viewed it as marginally more critical to be influenced by the anchoring effect if confidence was reduced before [*Ms* = 6.38 and 5.83, *SDs* = 1.21 and 1.63 with and without confidence reduction on a nine point scale, $t(1, 163) = 1.70$, $p < .09$]. Different from Study 2, participants did not consider the anchoring forewarning more while forming their own estimates after the confidence reduction.[8]

As expected, an index of the self-reported confidence measures which were assessed after the second task and at the end of the study (judgment certainty, self-perceived competence, judgment quality of the participant's own estimates compared to others: Cronbach's $\alpha = .71$) was significantly reduced by the confidence reduction [*Ms* = -0.15 and 0.19 (z-values) with and without confidence reduction respectively], $t(1, 163) = 3.21$, $p < .01$. In line with Studies 1 and 2, all other self-report measures (the probability of a hidden task difficulty, mood, frustration) were not influenced by the confidence reduction, $ts < 1.1$.

Magnitude of the anchoring effect

The same procedures of outlier exclusion, z-standardization, and reverse-scoring of low anchor tasks as before were used. Again, the two combined estimations for each expertise level were submitted as repeated measures to a 2 (expertise: high vs. low) x 2 (confidence reduction: reduction vs. no reduction) x 2 (forewarning: forewarning vs. no forewarning) ANOVA with expertise as a within-factor and confidence reduction and forewarning as between-factors.

[8] The expected trend was observable, but far from being significant (p=0.22).

Across all conditions, the anchoring effect was significantly different from zero (M = 0.52), $F(1, 139) = 232.99$, $p < .001$, $\eta^2_p = .59$. As in Study 2, a significant main effect was found for the anchoring forewarning. Anchoring effects were reduced by the forewarning [Ms = 0.41 and 0.62 with and without anchoring forewarning, $F(1, 162) = 8.83$, $p < .01$, $\eta^2_p = .05$]. Unlike in Study 2, the ANOVA revealed a significant main effect of expertise. The anchoring effect was weaker in the high expertise tasks (M = 0.40) than in the low expertise tasks (M = 0.62), $F(1, 162) = 9.35$, $p < .01$, $\eta^2_p = .06$.

The expected three-way interaction of expertise x confidence reduction x anchoring forewarning was also significant, $F(1, 162) = 4.77$, $p = .03$, $\eta^2_p = .03$, see Figure 4. Additionally, the confidence reduction x anchoring forewarning two-way interaction was marginally significant, $F(1, 162) = 3.75$, $p = .05$, $\eta^2_p = .02$. None of the remaining effects obtained significance, $Fs < 1$.

Magnitude of the anchoring effect: low expertise

In the low expertise estimation tasks, I predicted the anchoring forewarning to reduce the anchoring effect irrespective of the confidence manipulation. Accordingly, I found a significant main effect of the anchoring forewarning in the low expertise condition, $F(1, 162) = 5.43$, $p = .02$, $\eta^2_p = .03$, see Figure 4. When preceded by a forewarning, anchoring in low expertise tasks was weaker (M = 0.51) than without a forewarning (M = 0.73). As expected, the two-way interaction between anchoring forewarning and confidence reduction and the main effect of the confidence reduction were not significant for low expertise tasks, $Fs < 1$.

Magnitude of the anchoring effect: high expertise

In the high expertise estimation task, the anchoring forewarning was supposed to reduce the anchoring effect only if confidence had been weakened before. In line with this prediction, a two-way interaction between confidence reduction and anchoring forewarning was found, $F(1, 162) = 7.92$, $p < .01$, $\eta^2_p = .05$, see Figure 4. Single contrasts revealed that the anchoring forewarning failed to reduce the anchoring effect in the high expertise tasks without a preceding confidence reduction ($F < 1$), whereas it was significantly reduced by an anchoring forewarning after a confidence reduction, $F(1, 162) = 11.95$, $p < .001$, $\eta^2_p = .07$.

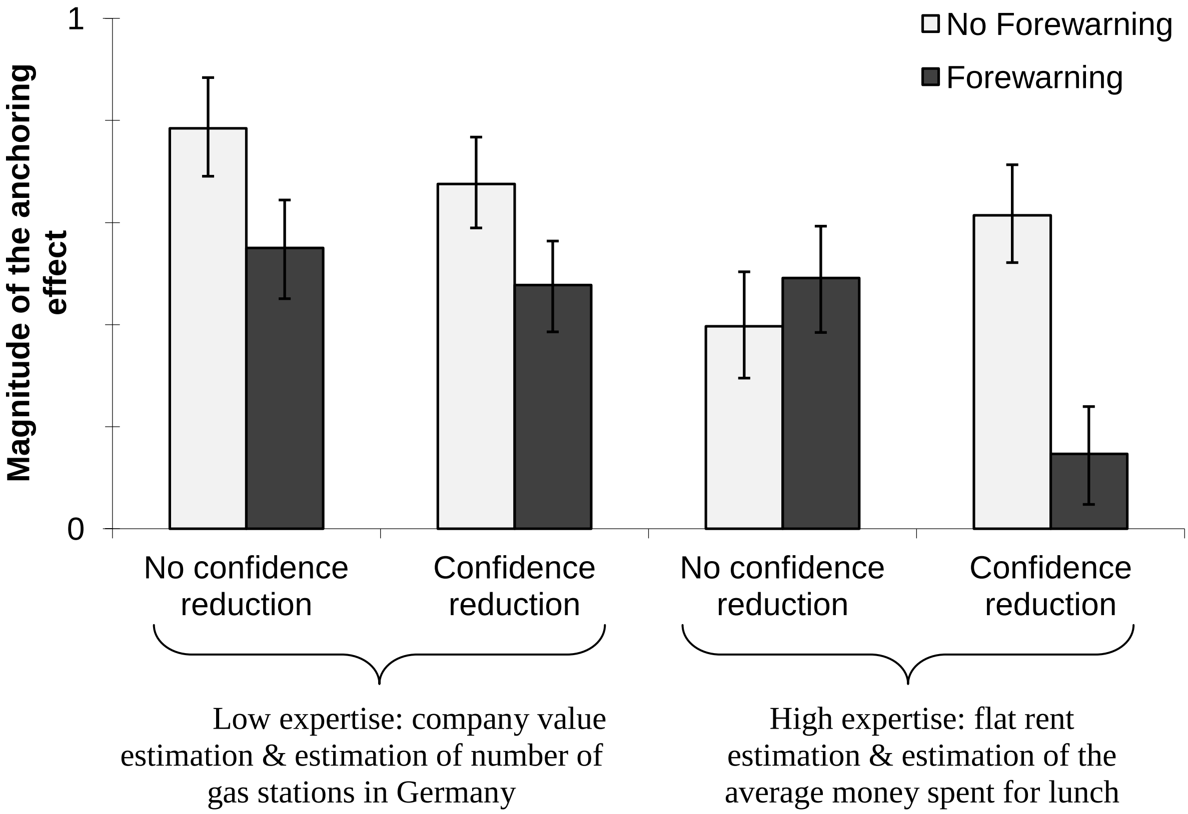

Figure 4: Magnitude of the anchoring effect (combined z-values with reverse coded low anchor items) by expertise (low vs. high), confidence reduction (no reduction vs. reduction), and forewarning (no forewarning vs. forewarning). Error bars represent standard errors. (Study 3)

In the high expertise task, the combination of the confidence reduction and the forewarning led to an anchoring effect that did not differ significantly from zero, $t < 1.5$. In all other conditions, the anchoring effect was significantly above zero, all $ts > 3.80$, all $ps < .001$.

Discussion

As in Study 2, key results are consistent with the assumption that expertise exerts a negative influence on the effectiveness of an anchoring forewarning due to experts' excessive confidence. In the low expertise tasks, the anchoring effect is reduced by the anchoring forewarning while the confidence reduction does not have any supplementary effect. In the high expertise tasks, the anchoring effect is affected by the forewarning only if confidence is first reduced.

In Study 3, the anchoring effect was significantly weaker in the high expertise condition than in the low expertise condition. This result deviates from Study 2, but also from other results in the literature which showed similar anchoring effects for different expertise levels (Englich et al., 2006; Northcraft & Neale, 1987). However, the result of Study 3 is not the first showing a reduction of anchoring effects in a setting with higher expertise (e.g., Smith, 2011; Wilson et al., 1996). In view of the deviating results and argumentations, it seems likely that the effects of expertise on the magnitude of anchoring effects may vary due to specific features of the anchoring task (e.g., the amount of anchor-consistent and anchor-inconsistent information available). To examine the influence of expertise on the anchoring effect independent of specific tasks, it seems promising to disentangle the confounding factors of expertise and the task itself by a between-subjects manipulation of expertise in Study 4.

Study 4

In the preceding studies, expertise was operationalized as experience and varied by the content of the estimation tasks, which were more or less familiar to the participants. On the one hand, this approach avoids some possible confounds which may occur if expertise is implemented by the often used method of comparing experts with non-experts (e.g., different age means of the two groups, different educational levels). On the other hand, it seems particularly important to examine the anchoring effect independent of specific tasks. Otherwise one cannot determine whether the weaker anchoring effect in the high expertise condition, which occurred in Study 3, is indeed caused by expertise itself, or by other characteristics of the estimation tasks. Additionally, one cannot be entirely certain that experience is an appropriate proxy for expertise in this context. It seems possible that only partial experience and “half knowledge” hinders the correction of the anchoring bias, whereas profound expertise does not hinder, but facilitates correction.

To extend the previous findings to the domain of real experts, I implemented a between-subjects manipulation of expertise in Study 4. In this study I compared a group of economic experts to a group of non-experts. Management consultants constitute a professional group that is often occupied with decisions and estimations in the world of economics. I therefore expected them to dispose of more expertise than students and former students in economic estimation tasks. As in Studies 1 to 3, this higher expertise should increase their confidence and undermine the beneficial effect of an anchoring forewarning, if no confidence reduction procedure is applied prior to the forewarning.

To verify my hypotheses, a 2 (expertise: high vs. low) x 2 (confidence reduction: reduction vs. no reduction) x 2 (forewarning: forewarning vs. no forewarning) x 2 (anchor: low vs. high) between design is applied.

Method

Procedure

Participants completed an online questionnaire designed with Unipark EFS Survey. The procedure was similar to the above-described studies. Again, participants were randomly assigned to one of the eight conditions and the order of the estimation tasks was counterbalanced across participants.

Materials

Expertise. Unlike in the preceding studies, Study 4 manipulated expertise by the professional economic knowledge of two groups of participants. All participants therefore completed the same four tasks which contained estimations of economic quantities. Two estimation tasks were identical to Study 3: (1) An estimation of a company value and (2) an estimation of the number of gas stations in Germany. In the third estimation task, the most important characteristics of a company (e.g., profit, industry cluster) were described. Participants then estimated the turnover of this company. As operating with variables like turnover or profit is part of a management consultant's daily business, their expertise was expected to be high. In the fourth estimation task, participants estimated the accumulated market shares of Mercedes, BMW, and Audi on the German automotive market. Due to the fact that the automotive sector is very important for the German economy, and that management consultants are often faced with the calculation and estimation of market shares, they are expected to be experts in this task, too.

In contrast, students and ex-students of disciplines other than economics and financial mathematics (who were excluded from participation) should not be familiar with such specialized issues (company turnover, market share, etc.) and their expertise should therefore be low.

Confidence reduction and anchoring forewarning. Confidence reduction and anchoring forewarning were identical to Study 3.

Anchor. As in Studies 1 to 3, participants first answered a comparative question and then provided an absolute estimate. In all groups, two of the tasks included a high anchor (100 billion € for the estimation of a company value, 550,000 for the number of gas stations, 16 billion € for the turnover estimation, and 70 % for the market share), the other two a low anchor (4 billion € for the estimation of a company value, 5,000 for the number of gas stations, 1 billion € for the turnover estimation, and 30 % for the market share). The order of the tasks was randomized. The high and low anchor values were set at the 85th percentile and at the 15th percentile on the basis of different pretests (n = 65 for the estimation of a company value, n = 34 for the number of gas stations, and n = 23 for the turnover and market share estimations).

Additional measures. As time is a scarce resource for management consultants, and redundant questions would possibly have increased dropout rates, the number of questions in the first part of the questionnaire was further reduced. Therefore, participants answered only one question about the subjectively experienced judgment certainty after each estimation task, skipping the blocked questions after the second item. Since dropouts are less critical after the key dependent measures, I added some questions at the last part of the study. Here, I mainly assessed the extent to which the confidence reduction changed in how far the anchoring forewarning was used by the participants because of their increased awareness about their own susceptibility. Participants therefore answered the same five questions about the anchoring forewarning and the awareness about the anchoring effect as in Studies 1 to 3. In order to measure their confidence, participants rated their general aptitude in estimation tasks, their performance in such estimation tasks compared to others, and the probability of a hidden task difficulty. Additionally, they indicated their general competence and expertise concerning estimation tasks and their general ability to make exact estimations. Finally, like in Studies 1 to 3, participants were inquired about their mood and frustration, answered questions about specific incidents during the experiment, and gave their demographic data, followed by an assessment of their effort to give good estimates.

Participants

For the non-expert sample, 95 students and former students (68 female, $M_{\text{Age}} = 23.95$ years, $SD_{\text{Age}} = 6.37$ years) recruited from a pool for online data-collection were given the chance to win a gift voucher by enrolling in a lottery. To ensure a low expertise level concerning economic tasks, students of economics and financial mathematics were kept from participating. For the expert sample, 64 management consultants (13 female [9], M_{Age} = 26.7 years, SD_{Age} = 9.02 years, mean employment time in economics = 3.6 years) were recruited via e-mail. Participants were randomly assigned to one of the experimental conditions. Due to implausibly long completion times, repeated participations, and clearly deficient understanding of the instructions, 12 participants (8 %) were excluded from further analyses.

Results

Manipulation check

Expertise. As expected, management consultants felt significantly more certain (M = 5.13, SD = 1.10 on a nine point scale) than non-experts (M = 3.35, SD = 1.20), $t\,(1, 138) = 11.93$, $p < .001$].

Confidence reduction. As the seven different confidence measures were highly correlated (perceived judgment certainty after each estimation task[10], participants' general aptitude in estimation tasks, their performance in such estimation tasks in comparison to other participants, the probability of a hidden task difficulty, their general competence and expertise concerning estimation tasks, and their general ability to give exact estimations: Cronbach's α = .81), I calculated one aggregate confidence index. This index was

[9] Due to the fact that the mean age and the male-to-female ratio differed significantly between the two groups, I alternatively conducted the main analyses with age and gender as covariates. This did not change the pattern of results, so that these analyses are not reported here.

[10] Different from Study 1 to 3, it was possible to include perceived judgment certainty after the estimation tasks in a confidence index, because expertise was manipulated between subjects.

marginally lower with the confidence reduction (*M* = -0.06, *SD* = 0.71) than without the confidence reduction (*M* = 0.08, *SD* = 0.82), $F(1, 135) = 2.77$, $p < .10$, $\eta^2_p = .03$.

Correspondingly, personal influencability by the anchoring effect was also rated as marginally higher after the confidence reduction (*M* = 6.14, *SD* = 1.43) than without the confidence reduction (*M* = 5.58, *SD* = 1.41) on a nine point scale, $F(1, 135) = 3.86$, $p = .05$. Unlike in Study 3, participants did not rate it more critical to be influenced by the anchoring effect with the confidence reduction than without the confidence reduction.[11] These inconsistent findings may possibly be explained by the position at the end of the questionnaire, which could result in increased variance because of tired participants and the delay between the confidence reduction treatment and the assessment of the dependent variables.

All other self-report measures (mood, alertness, consideration of the anchoring forewarning while forming estimates, effort to provide good estimates) did not yield any significant effects, all *ts* < 1.02, *ps* > 0.33.

Magnitude of the anchoring effect

The same procedures of outlier exclusion, z-standardization, and reverse-scoring of low anchor tasks as in Study 1 and 2 were used. The combined estimations of the four tasks of one expertise level were then submitted to a 2 (expertise: high, management consultants vs. low, students) x 2 (confidence reduction: reduction vs. no reduction) x 2

[11] The expected pattern was observable, but far from being significant (p=0.28).

(forewarning: forewarning vs. no forewarning) ANOVA with expertise, confidence reduction, and forewarning as between-factors.

Across all conditions, the anchoring effect was significantly higher than zero (M = 0.53), $F(1, 139) = 224.89$, $p < .001$, $\eta^2_p = .62$. As in Studies 2 and 3, the ANOVA revealed a significant between-subjects effect for the anchoring forewarning. Anchoring effects across both expertise levels were stronger without a forewarning [Ms = 0.42 and 0.64 with and without anchoring forewarning, $F(1, 139) = 12.24$, $p < .001$, $\eta^2_p = .08$]. Similar to Study 3, but different from Study 2 a significant main effect of expertise appeared. The anchoring effect was weaker in the high expertise tasks (M = 0.38) than in the low expertise tasks (M = 0.64), $F(1, 139) = 16.34$, $p < .001$, $\eta^2_p = .10$. Apparently, the anchoring effect affected the management consultants to a lesser degree than student subjects. The expected expertise x confidence reduction x anchoring forewarning three-way interaction was also significant, $F(1, 139) = 4.81$, $p = .03$, $\eta^2_p = .03$. To better understand this complex interaction, the combined estimations of the four tasks of one expertise group (students and management consultants) were separately submitted to a 2 (confidence reduction: reduction vs. no reduction) x 2 (forewarning: forewarning vs. no forewarning) ANOVA.

Magnitude of the anchoring effect: low expertise (students)

In the low expertise group, the anchoring forewarning is expected to reduce the anchoring effect irrespective of the confidence reduction. Correspondingly, results show a significant main effect of the anchoring forewarning, $F(1, 80) = 8.77$, $p < .01$, $\eta^2_p = .10$, see Figure 5. When preceded by a forewarning, anchoring was weaker (M = 0.78) than without a forewarning (0.51). As expected, the two-way interaction between confidence reduction and anchoring forewarning was not significant in the low expertise group, $F < 0.01$. Finally, there was no main effect of the confidence reduction, $F < 1.4$.

Magnitude of the anchoring effect: high expertise (consultants)

In the high expertise condition, the anchoring forewarning was supposed to reduce the anchoring effect only if confidence had been reduced prior to the forewarning. Accordingly, results show the expected two-way interaction between confidence reduction and anchoring forewarning, $F(1, 59) = 4.1$, $p = .05$, $\eta^2_p = .06$, see Figure 5. Additionally, the ANOVA reveals a main effects for the anchoring forewarning [$F(1, 59) = 4.35$, $p < .04$, $\eta^2_p = .07$] and the confidence reduction [$F(1, 59) = 4.76$, $p < .03$, $\eta^2_p = .08$]. Single contrasts show that the anchoring forewarning failed to reduce the anchoring effect

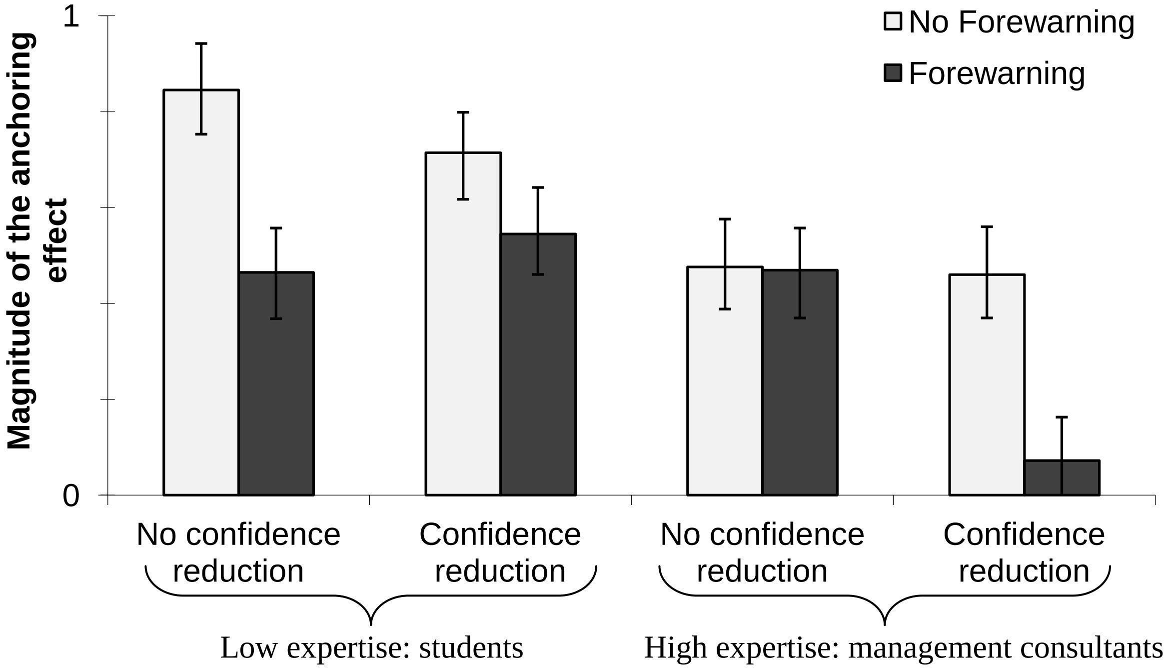

Figure 5: Magnitude of the anchoring effect (combined z-values with reverse coded low anchor items) by expertise (low vs. high), confidence reduction (no reduction vs. reduction), and forewarning (no forewarning vs. forewarning). Error bars represent standard errors. (Study 4).

without a preceding confidence reduction [Contrast anchoring forewarning (forewarning vs. no forewarning) *without* confidence reduction: $F(1, 59) = 0.16$, $p = .90$] in the high expertise group. However, the forewarning significantly reduced the anchoring effect after the confidence reduction [Contrast anchoring forewarning (forewarning vs. no forewarning) *with* confidence reduction: $F(1, 59) = 7.93$, $p < .01$, $\eta^2_p = .12$].

In the high expertise group, the combination of the confidence reduction and the forewarning led to an anchoring effect that did not differ significantly from zero, $t < 1$. In all other conditions, the anchoring effect was significantly higher than zero, all $ts > 8.55$, all $ps < .001$.

Mediation analysis

The results of the main analysis suggest that in the case of management consultants (experts), the effect of a confidence reduction on the impact of a forewarning relied on confidence ratings. To formally test this mediation, I followed the bootstrapping procedure of Preacher and Hayes (2008), computing a confidence interval for the indirect effect (the path including the mediator, see Figure 6). If zero falls outside this interval, mediation will be present. To determine the confidence interval, I utilized the SPSS (Statistical Package for the Social Sciences) macros of Preacher and Hayes (2012). In this analysis, confidence reduction was the independent variable, while the magnitude of the anchoring effect was the dependent variable. As the seven aggregated confidence measures and the three measures assessing the awareness about the anchoring effect (personal influencability by the anchoring effect, attitude towards being influenced by the anchoring effect, consideration of the anchoring forewarning while forming estimates) were significantly correlated, *r* (59) = .28, *p* = .03, I used an index of these ten measures (Cronbach's α = .85) as a mediator. The mediator, therefore, does not only represent confidence levels, but also comprises awareness of the anchoring effect. Results of this procedure revealed a 95% confidence interval ranging from −0.40 to −0.007. The fact that zero did not fall inside the limits of this interval indicates a mediation effect. This finding supports the view that the confidence reduction in a high expertise setting allows for a beneficial effect of the anchoring forewarning.

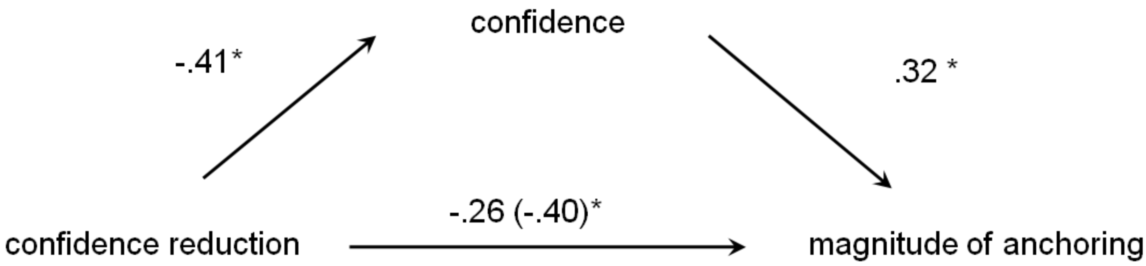

Figure 6: Path coefficients for mediation in Study 4. The coefficient in parentheses represents the direct effect of confidence reduction on the magnitude of the anchoring effect. + $p < .10$, * $p < .05$, ** $p < .01$.

Discussion

Like in Studies 1 to 3, results confirm my assumption that expertise exerts a negative influence on the effect of an anchoring forewarning because of experts' excessive confidence. Non-experts were able to correct for the bias after the forewarning irrespective of the confidence reduction. Experts, on the other hand, only profited from a forewarning if their initial confidence was reduced. In this condition, the anchoring effect was not statistically significant anymore.

Self-report ratings give further support for my reasoning. As in Studies 2 and 3, confidence ratings were significantly elevated in the high expertise condition. The confidence reduction significantly lowered these confidence ratings. The important influence of confidence on the impact of anchoring forewarnings was also apparent in the mediation analysis. Results demonstrated that the amount of confidence and awareness about the anchoring effect is the mediator driving the success of an anchoring forewarning. Further evidence comes from participants rating their own influenceability by the anchoring effect higher after the confidence reduction. This corroborates the assumption that people may be more aware about their own susceptibility to the anchoring effect after a confidence reduction.

GENERAL DISCUSSION

Significance of the present findings

Common sense makes us assume that experts should be less prone to the risks of irrational biases. As explained previously, there are also more substantial reasons to believe that experts should be better able to correct for judgmental biases such as the anchoring effect than non-experts. First, experts should correct more efficiently because they have more anchor-inconsistent information available (c.f., Englich, 2008), which can reduce anchoring effects (Mussweiler et al., 2000). Second, higher levels of expertise limit the range of responses considered as plausible and therefore may reduce the anchoring effect. Third, experts may also be more aware of the direction and magnitude of a potential bias due to their broader and more profound knowledge in their specific domain of expertise.

As laid out before, experts are generally more overconfident than non-experts (e.g., Englich et al., 2006; Törngren & Montgomery, 2004). Such heightened confidence might hinder experts from listening carefully to forewarnings on their own influenceability, such as, forewarnings about anchoring effects (Wilson et al., 1996). In a similar vein, people with higher expertise tend to neglect helpful decision aids (Arkes et al., 1986). Thus, excessively confident experts may fail to correct their estimates following an anchoring forewarning.

In line with this reasoning, the results of four successive studies demonstrated that an anchoring forewarning merely consisting of information on the existence and robustness of the anchoring effect can be a successful correction strategy for the specifically robust anchoring effects of externally provided numerical standards. Even more importantly, they repeatedly demonstrate that for experts, a confidence reduction is a necessary prerequisite to enable them to benefit from an anchoring forewarning (Studies 2 to 4). If experts' confidence in their own decisions was not reduced, experts did not benefit from an anchoring forewarning, whereas non-experts did gain such benefit. Additionally, for management consultants who received a forewarning, an aggregate of confidence and awareness mediated the effect of the confidence reduction on the magnitude of anchoring (Study 4). This finding supports our reasoning about heightened confidence as the reason for the failure of the anchoring forewarning in reducing anchoring in the high expertise conditions.

My findings offer several key contributions to the judgmental anchoring, expertise, and overconfidence literature. For the first time, a successful correction of experimenter-provided anchors by a general debiasing method could be shown, and replicated, in a standard anchoring paradigm. In self-generated anchor paradigms, the anchoring effect has repeatedly been reduced by measures such as forewarnings (Epley & Gilovich, 2005) or incentives (Epley & Gilovich, 2005). In contrast, in an experimenter-provided standard anchoring paradigm, general debiasing methods such as forewarnings or incentives have been repeatedly shown to fail in reducing anchoring (Epley & Gilovich, 2005; Wilson et al., 1996). Only specific strategies like considering-the-opposite (Mussweiler et al., 2000) or incentives under the condition of the direction of adjustment being evident (Simmons et al., 2010) have been successful. However, in these studies, it was ensured either that anchor-inconsistent knowledge was activated (Blankenship et al., 2008; Galinsky & Mussweiler, 2001; Mussweiler et al., 2000) or that participants knew about the direction of the adjustment (Simmons et al., 2010). Consequently, the current research is the first to demonstrate the correction of the anchoring effect in a standard anchoring paradigm without support of the experimenter. This finding is particularly important for applied contexts, in which it is often unrealistic to give support that is specific to a certain task.

Second, the here presented studies are the first research to systematically examine the impact of expertise on the effectiveness of a forewarning. Although previous research has consistently supported the vulnerability of experts to anchoring effects (Englich et al., 2006; Northcraft & Neale, 1987) as well as to other decision biases (e.g., Guthrie, Rachlinski, & Wistrich, 2001), it has not previously been shown how expertise affects the different stages of the debiasing process. My results demonstrate that expertise can be a double-edged sword for the risk of falling prey to the anchoring effect. On the one hand, the anchoring effect was weaker in the high expertise condition (Studies 3 & 4). This suggests that higher expertise can have direct beneficial effects on the susceptibility to anchoring effects. On the other hand, my results show that experts did not benefit from anchoring forewarnings without an additional reduction of confidence. This ambivalence might hold true for other biases as well.

Third, my results advance our understanding of why forewarnings and other debiasing attempts are often so surprisingly ineffective. As far as my research suggests, this is due to heightened confidence, which may prevent people from using information given in the forewarning. Although I only show this for the domain of anchoring effects, examining whether this also holds for other biases would appear to be a promising approach, because heightened confidence is likely to affect the openness to bias forewarnings in general.

In the following, I first discuss a possible alternative explanation of my findings. Afterwards, I elaborate on the relation of my results to previous research in this field. In the then following part, I propose directions for future research. Finally, I suggest practical implications resulting from my findings.

Alternative explanation: Intentional usage of the anchor value

Subjects who anticipate the experimenter to be maximally informative may infer that the provided anchor value is close to the true value (Jacowitz & Kahnemann, 1995; Schwarz, 1994). Non-experts may be more open to seeing the randomly generated anchor value as informative, because they possibly simply do not know the answer and therefore intentionally use the anchor value as information. This process could explain why anchoring was stronger for non-experts than for experts in Studies 3 and 4. If the anchor is intentionally used as information, this may add to the other mechanisms of anchoring and cause a stronger anchoring effect in the low expertise condition.

Moreover, if information is used intentionally, one should be more aware of the influence of information. Awareness makes it easier to correct an influence (Strack & Hannover, 1996). It is hence plausible that preventing the intentional usage of the anchor is easier than reducing the amount of the anchoring effect caused by unintentional processes (numerical anchoring, insufficient adjustment, selective activation of information, or scale distortion). Based on this reasoning, one would predict the pattern which occurred in the conditions without confidence reduction: The forewarning should be more effective for non-experts than for experts, because a larger proportion of non-experts' anchoring effect could be caused by intentional usage of the anchor and could therefore be easier to correct.

However, in the confidence reduction conditions, a forewarning successfully reduced experts' anchoring effects in my studies, whereas it did not have a beneficial impact for non-experts. If the success of the forewarning was only driven by the amount of intentional usage of the anchor, the success in the high expertise condition necessarily has to be explained by less intentional usage of the anchor after the confidence reduction. The reduction procedure in Studies 3 and 4 is indeed likely to make subjects alert to the risk of using irrelevant information which seems relevant at first glance. But, this reasoning cannot explain why the confidence reduction had no effect in the low expertise condition. In contrast, in the low expertise condition, one would expect a stronger effect of the confidence reduction because of more intentional usage of information without confidence reduction. Intentional usage of information can therefore not explain why a confidence reduction did not have any effect on the magnitude of non-experts' anchoring effect.

Relation to previous research

Compatibility with previous research showing a failure of forewarnings

At first glance, one may assume that my findings contradict the results of Wilson et al. (1996) as well as Epley and Gilovich (2005). Here, a forewarning failed to reduce anchoring in an experimenter-provided anchoring paradigm even with a non-expert sample. However, this discrepancy can possibly be explained by the lower expertise level of my low expertise tasks in contrast to those used by Wilson et al. (1996) and Epley and Gilovich (2005). In these studies, expertise levels were not varied. In the study by Wilson et al. (1996), the only task consisted of estimating the number of doctors in the telephone book of the participants' home town; in the study by Epley and Gilovich (2005), participants answered general knowledge questions such as "What is the population of Chicago?". These questions are closer to participants' daily experiences than an estimation of a company's value, because it is likely that participants have looked up doctors in the telephone book before, and have probably also received information about the population of Chicago at some point. Therefore, my low expertise tasks (e.g., the estimation of a proper company value) represent an even lower expertise level. If participants have very low expertise for an estimation task, forewarnings might be sufficient in reducing the bias.

The effect of expertise on the magnitude of the anchoring effect

In Studies 3 and 4 of this dissertation anchoring was less pronounced in the high expertise setting than in the low expertise setting. To date, a less pronounced anchoring effect for higher expertise or more knowledge has often been shown for basic anchoring (e.g., Englich, 2008; Kaustia et al., 2008; Wilson et al., 1996). In standard anchoring paradigms, Blankenship et al. (2008) and Mussweiler et al. (2000) showed that anchoring was reduced if anchor-inconsistent knowledge was administered before the anchoring task. Nevertheless, the purposeful administration of anchor-inconsistent knowledge clearly differs from the comparison of high expertise and low expertise in my studies. A successful reduction of the anchoring effect by learning relevant knowledge has recently been demonstrated even though this knowledge was not necessarily inconsistent with the anchor values (Smith, 2011, Studies 1 to 3). However, respective findings differ. Englich (2008) did not find a reduction of standard anchoring by providing knowledge prior to the anchoring tasks. Even more importantly, in my studies, knowledge concerning the task was not learned during the experiment but differed from the beginning. In other research in which initial levels of expertise differed, this did not influence the magnitude of anchoring (Englich et al., 2006; Northcraft & Neale, 1987; Smith, 2011, Study 4). Taken together, most of the research with similar methods as in this dissertation finds no effect of expertise on the magnitude of anchoring, whereas I found such an effect in two of the four studies.

An explanation for the deviating results of the latter studies and my own research might lie in the extent to which expertise varies. In support of this reasoning, an inverted U-shaped relationship between expertise and hindsight bias has been proposed (Knoll, 2010). More specifically, hindsight bias was exacerbated by expertise up to a certain level. After reaching this level hindsight bias was reduced by expertise. If a similar pattern also occurs for the anchoring effect, this can result in different effects of a variation of expertise depending on the absolute level of expertise. For example, Englich et al. (2006) compared experts in criminal law and experts in other fields of law concerning the strength of the anchoring effect in a task in which participants estimated an appropriate sentencing demand in a shoplifting case. Although knowledge of experts in criminal law is clearly superior, experts in other domains of law were also somewhat familiar with criminal cases because they are lawyers. Comparably, the expert sample recruited by Northcraft and Neale (1987), namely real estate agents, were clearly competent in estimating the value of real estate. But, in contrast to my research, at least a proportion of their student reference group was probably familiar with real estate transactions, too. Accordingly, 14.6% of the students reported that they had already been involved in a real estate transaction. As their average age was twenty-two years, and they had probably lived at their parents' homes for most of their lives, it seems likely that an even higher percentage were informed about the prices of real estate transactions without direct involvement. Taken together, in both studies cited, experts were compared with "semi-experts", which might have caused the identical anchoring effects across groups.

By contrast, an intended goal of my research was to vary expertise to a large extent in order to study the moderation of expertise levels on the effects of an anchoring forewarning. Therefore, it is far less likely that students in my studies were familiar with the included topics (e.g., estimation of a company value or turnover estimations), whereas management consultants are very familiar with these topics. Possibly, only this pronounced difference of expertise in conjunction with the consequent high statistical power is able to show the effects of expertise on the strength of the anchoring effect. Lower statistical power could also explain why these effects did not occur in Studies 1 and 2, because they involved fewer participants and fewer estimation tasks.

Furthermore, a different level of subjectivity could explain the deviating results of my own research and the studies of Englich et al. (2006) and Northcraft and Neale (1987). Whereas my tasks (e.g., the calculation of an adequate value of a company) require more conceptual problem-solving, such as rough estimations, a sentencing requirement is more dependent on the subjective attitudes of a judge (Hogarth, 1971; Partridge & Eldridge, 1974). Although this contrast is probably less pronounced between my tasks and the estimation of the value of real estate, the latter is in a way still dependent on subjective preferences (one person prefers to live downtown, while another person prefers a calm location in a suburb). Possibly, this subjectivity of the estimations allows more selective activation of available anecdotal memory content (which experts possess in particular) and therefore undermines the potentially positive effects of expertise (e.g., a narrower interval of plausible responses, more anchor-inconsistent knowledge).

Implications for the debate on the mechanisms of anchoring

The insufficient adjustment account predicts that the magnitude of the observed anchoring effect should vary with the extent of effortful thinking devoted to a task (Epley & Gilovich, 2001, 2005, 2006). The extent of effortful thinking is most typically manipulated by incentives (Epley & Gilovich, 2005; Simmons et al., 2010; Wilson et al. 1996). But forewarnings are also seen as manipulations that trigger effortful thinking

(Chapman & Johnson, 2002; Epley & Gilovich, 2005). According to this reasoning, findings demonstrating the ineffectiveness of forewarnings in reducing the effect of externally provided anchors (Epley & Gilovich, 2005; Wilson et al., 1996) are often interpreted as empirical evidence against the insufficient adjustment account (Chapman & Johnson, 2002; Epley & Gilovich, 2005).

Different from this previous empirical evidence, Studies 1 to 4 of the here presented dissertation show that forewarnings are effective in reducing the effect of externally provided anchors under certain conditions. One might argue that this finding may speak in favor of the insufficient adjustment account, which predicts that forewarnings should affect the magnitude of the anchoring effect. Conversely, it may be an argument against other approaches (e.g., the selective accessibility model) that contradict the effect of accuracy motivation on the magnitude of anchoring (Chapman & Johnson, 2002; Simmons et al., 2010).

However, the different explanations of the anchoring effect are mostly not seen as mutually exclusive, but as complementary (Epley & Gilovich, 2010; Mussweiler, 1997; Mussweiler & Strack, 1999, 2000; Mussweiler et al., 2004; Simmons et al., 2010; Wegener et al., 2010a, Wilson et al., 1996). Consistently, the success of forewarnings to reduce anchoring in my studies is probably not a meaningful argument in favor of the insufficient adjustment account or against other accounts. Rather, it might show that the insufficient adjustment account is able to explain a relatively large proportion of the anchoring effect in these specific studies. In other studies with a different operationalization, this might be different.

Implications for the revised theory of anchoring and adjustment

The revised theory of anchoring and adjustment attempts to overcome the distinction between self-generated and externally provided anchors (Simmons et al., 2010). The failure to reduce the effect of externally provided anchors by incentives and forewarnings is explained by the fact that people do not know about the correct direction of adjustment. If they did, an increase of accuracy motivation should reduce the effect of provided anchors (Simmons et al., 2000). This has been demonstrated, for example, by showing that incentives reduced the anchoring effect of implausible anchors, whereas they did not reduce the anchoring effect of plausible anchors (Simmons et al., 2010, Studies 3a & 3b). As people should be more certain about the correct direction of adjustment if anchors are implausible, this is seen as empirical evidence in favor of the revised theory of anchoring an adjustment.

In Study 4 of this dissertation, the effectiveness of a forewarning in reducing the effect of provided anchors in economic estimation tasks was compared between non-experts and experts. For experts, the same anchor value should be less plausible than for non-experts, because the former should have more experience and rules of thumb available to narrow the range of potential answers. As non-experts and experts received the same anchors, one would expect that experts should be more certain about the direction of adjustment. Because of this higher certainty, the revised theory of anchoring and adjustment (Simmons et al., 2010) would predict a forewarning to be more effective in the high expertise task than in the low expertise task. However, my results demonstrate exactly the opposite pattern: The forewarning was less successful in reducing the anchoring effect in the high expertise task than in the low expertise task. My results are therefore disparate to the revised theory of anchoring and adjustment. Possibly, this contradiction can be explained by the use of different methods in the Simmons et al. (2010) and my studies. Simmons et al. (2010) used a two-stage process. The first stage consisted of anchoring tasks similar to those used in my studies. In the second stage, the participants were informed about the possibility to revise their estimates and that they could earn points if their estimate was close to the true value. This type of instructions may prompt participants to revise their estimates, because they intuitively consider it a hint that their initial number is wrong. They may therefore intuitively modify their estimates, even if they do not know about the correct direction of their revision. Such an intuitive modification in the correct direction but equally often in the incorrect direction, should not reduce the anchoring effect. Supporting this reasoning, Simmons et al. (2010) could show that a reduction of the anchoring effect only occurs in this paradigm, if it is ensured that participants know the correct direction of an adjustment.

In contrast, the forewarning in my studies just included information about the direction of the anchoring effect, but did not include a request to correct. It is therefore less probable that participants, although ignorant of the direction of adjustment, intuitively modified their values. Taken together, the effects observed by Simmons et al. (2010) may be specific to their method, which may prompt participants to revise their estimates without knowing the correct direction of adjustment.

Future directions

Effect of expertise and confidence on forewarnings about other biases

As overconfidence is a very widespread phenomenon (Moore & Swift, 2011; Plous, 1993), and there is no apparent reason to believe that my findings are unique to the domain of anchoring, they may have important implications for other biases as well (e.g., hindsight bias, confirmation bias). More precisely, it seems likely that my findings on the influence of expertise and confidence on the effectiveness of forewarnings also apply to other biases. Similar processes may make people refrain from correcting their judgments after a forewarning about other biases. They may believe that only people in general are susceptible to the respective bias, but neglect their own susceptibility. Consistent with this reasoning, people see the existence of many other biases much more in others than in themselves (Pronin & Kugler, 2007; Pronin, Gilovich, & Ross, 2004; Pronin, Lin, & Ross, 2002). It might therefore be promising to assess the impact that confidence and expertise have on the effectiveness of debiasing measures in the context of other biases.

High power as an additional source of overconfidence

Recently, it has been shown that the experience of high social power leads to more overconfident decision making (Fast, Sivanathan, Mayer, & Galinsky, 2012). Social power is the possibility to influence others (De Dreu & van Kleef, 2004). In research by Fast et al. (2012), the induction of high power, for instance by an episodic recall task, led to overconfident decisions with monetary losses for the powerful. In this process, objective power only produces overconfident decision making if it leads to a subjective feeling of power (Fast et al., 2012, Studies 4 & 5).

As I am able to show that heightened confidence exerts detrimental effects on the effectiveness of forewarnings, it seems plausible to assume that the experience of power (which leads to overconfidence) may also reduce the effectiveness of forewarnings. This may further exacerbate the risk that professional decision makers like politicians or managers will not correct their decisions after forewarnings, because they are not only experts but also experience high power. The experience of coming to an important decision affecting a large number of people may induce a feeling of power. If this reasoning holds true, this would lead to the dangerous tendency that the more important a decision is, the less decision makers will correct it for decision biases. It therefore seems important to examine whether higher power does, in fact, exert a detrimental influence on forewarnings in particular or debiasing measures in general.

Additionally, it has been shown that the experience of power elicited by power gestures increases confirmatory information processing (Fischer, Fischer, Englich, Aydin, & Frey, 2011). Confirmatory information processing comprises biased assimilation (Lord, Ross, & Lepper, 1979) and selective exposure (Festinger, 1957). Biased assimilation is defined as an overestimation of the relevance, quality, and importance of information which is consistent with one's own point of view, as opposed to inconsistent information (Ditto & Lopez, 1992; Lord et al. 1979). Selective exposure denotes people's systematic preference for information which is consistent with their point of view (Festinger, 1957; Fischer, Kastenmüller, Greitemeyer, Fischer, & Frey, 2011; Lundgren & Prislin, 1998).

Anchoring, too, is often explained by selective accessibility (Mussweiler & Strack, 1999, 2000a) or confirmatory search (Chapman & Johnson, 1999), according to which judges selectively search for and activate hypothesis-confirming knowledge about the judgmental target. This definition of the process reflects the parallels between anchoring and selective exposure. In both cases, consistent information is preferred – consistent with the anchor (anchoring) and with the judge's point of view of the (selective exposure) respectively. In view of these similarities, it should be examined whether the experience of power by power gestures also leads to more anchoring due to more anchor-consistent information processing.

Overconfident or merely more confident?

The term overconfidence implies an excess of confidence. In most of the research demonstrating this excess, subjective ratings of confidence are compared to a correct response (i.e., an objective or empirical result) (Pallier et al., 2002). For example, in order to measure overconfidence, people are asked to estimate how many correct answers they will give in a ten-item quiz (cf., Moore & Healy, 2008). This subjective rating of confidence is then compared to the actual number of correct answers (explicit measure). If people were objective about their skills, only random deviations between the subjective ratings and the correct response, but no systematic difference should occur (Pallier et al., 2002; Phillips, 1973). Such a systematic difference, however, can be found in a multitude of situations, reflecting people's tendency to generally overrate their performance (Moore & Swift, 2011; Pallier et al., 2002; see theory section for a more detailed overview).

In the four studies in this paper, participants were asked to give subjective ratings of confidence. For example, they were asked to evaluate their ability to answer such estimation tasks (Studies 2 to 4). However, there was no objective or empirical result implemented in my studies. It is therefore not possible to calculate the exact difference between subjective ratings and an objective or empirical result. In any case, results demonstrate that experts are more confident than non-experts. Additionally, a confidence reduction makes an anchoring forewarning more effective. Although I cannot clearly conclude the extent of overconfidence from these findings, they do provide an indication that heightened confidence of experts can be harmful and that a reduction can be helpful in reaching better decisions. Based on this reasoning, experts' unreduced confidence levels can be labelled overconfident because they led to inferior adjustment, even though no valid standard of confidence excess was present.

In order to measure whether experts are indeed overconfident and not merely more confident than laymen, future research should additionally include an objective or empirical measure.

Developing a better understanding of the underlying processes

This dissertation focuses mainly on the benefits of effective debiasing measures for the applied context. A significant contribution of future research might lie in achieving a better understanding of the underlying processes. Earlier, I proposed that experts' overconfidence undermines the beneficial effect of an anchoring forewarning, because overconfidence prevents them from noticing the bias at work. According to the model of mental contamination and mental correction (Wilson & Brekke, 1994, see Figure 1), awareness of the biasing influence is a necessary precondition for successful correction processes.

My studies mostly provide evidence for the first part of my process assumption, namely that the success of an anchoring forewarning is mediated by the level of confidence. First, the repeatedly significant confidence reduction x anchoring forewarning interaction in the high expertise settings demonstrates the impact of the confidence reduction, given that initial confidence is high. In the low expertise setting, confidence did not have such an effect. Additionally, the indirect effect in the bootstrapping analysis in Study 4 indicated that the effect of a confidence reduction on the impact of a forewarning was indeed mediated by confidence levels in the group of management consultants (experts). Both facts argue for the causal role of confidence in the failure of forewarnings.

However, my research provides limited insight of how the success of a forewarning depends on the awareness of the own susceptibility to the anchoring effect. Although the term awareness may implicate measurability by explicit questions, the major problem lies in the distinction between awareness of the anchoring bias and the level of confidence. For example, if an individual indicates not being aware of the anchoring bias, this may indeed reflect ignorance of the bias, but it might just as well be a sign of overconfidence. My empirical results confirm that it is difficult to measure confidence and awareness of the anchoring effect independently. For example, in Study 4, I found a significant correlation of confidence and anchoring awareness measures.

To attain a better understanding of the distinctive role of awareness of the anchoring effect, future research should strive to find a measure of awareneyy which is independent of the level of confidence.

Ensure long-term effects of a forewarning

For practical applications, for example in companies, it would be desirable if debiasing methods reduced biases not only directly after the usage of the method. On the contrary, such methods should be able to reduce biases sustainably, because undertaking a debiasing method before each potentially biasing influence seems unrealistic. A continuous monitoring of potentially biasing influences or the repeated application of the debiasing method by a company, for example, would probably be very costly with regard to time and organizational resources. Since, to date, a durably successful reduction of the anchoring effect has not been demonstrated, research should strive to find such methods.

In future research, it could therefore be promising to test if an anchoring forewarning continues to be successful after a delay. This should be tested independently for different expertise levels, because the higher confidence of experts could hinder the positive effects of the forewarning.

Additionally, the forewarning is a very short intervention, which is probably difficult to remember after a certain delay. An improved memorization of the anchoring forewarning could make the forewarning more successful after a delay. Memorization can be enhanced by repeated activation of this specific information (Anderson, 2000; Collins & Loftus, 1975) or repeated activation of information which is semantically related to this information (Anderson, 1983; Collins & Loftus, 1975). Consequently, to improve the memorization of the anchoring forewarning, one could ask participants to recall the forewarning after a certain delay, ask them to use the anchoring forewarning in other tasks, or train them to correct other decision biases which are semantically related to the anchoring forewarning. Memorization can also be improved by making processing more elaborate (Craik, 2002; Craik & Lockhart, 1972; Lockhart, 2002). In this sense, providing information about related decision biases or teaching participants about the theoretical background of anchoring also seem to be possible interventions to ameliorate the memorization of the anchoring effect. Future research should further investigate these options.

Practical implications

Confidence as a risk for experts' decision making

As decision makers like politicians or company executives are often experts in their fields, they may feel particularly confident. Moreover, leaders' historical success in their roles (Bassarab, 2011) and their high power (Fast et al., 2012) makes them particularly confident. As a consequence, these experts are less likely to correct their decisions due to forewarnings. The failure of forewarnings to reduce the effect of undesired influences is particularly dangerous because people often believe that they have the ability to resist biases by merely knowing about them (Wilson, Centerbar, & Brekke, 2002). Moreover, it is one of the commonly used debiasing methods to forewarn experts about biases in order to create awareness of their existence (Kaufmann et al., 2009). My studies point out that this common belief is often wrong. An anchoring forewarning is only effective in reducing the anchoring effect if expertise is low or confidence is reduced. As overconfidence may indeed be found in many leaders and decision makers of the financial market, the findings of my study suggest one possible explanation why the forewarnings of Roubini (2007) and others could not help to prevent the financial and economic crisis in 2007.

Confidence should be carefully reduced

A reduction of experts' confidence has been demonstrated to make anchoring forewarnings effective in the case of experts. Although heightened confidence obviously exerts negative influences in this case, one should not forget that although overconfidence may be a curse in this context, it can also be a blessing in other situations. For example, overconfidence of managers drives innovation (Galasso & Simcoe, 2011), overconfident self-perception can increase outcomes in a team setting (Ludwig, Wichardt, & Wickhorst, 2011), and managerial overconfidence and optimism can increase a company's value, because rational managers postpone the decision for longer than is in the best interest of shareholders (Gervais, Heaton, & Odean, 2002). Moreover, depressive realism stresses that depressed subjects are realistic and non-depressive subjects are overconfident (Alloy & Abramson, 1979, 1982; Vázquez, 1987; Von Helversen, Wilke, Johnson, Schmid, & Klapp, 2011). Although this finding clearly does not allow the causal conclusion that a reduction of overconfidence makes people feel depressed, the co-occurrence of depression and reduced overconfidence should at least make us alert regarding possible negative side effects of confidence-reducing measures. To control for negative side effects of the confidence-reducing measures in this dissertation, I assessed mood. In the here presented studies, the confidence reductions did not have any effects on mood.

Involving non-experts in important decisions

My findings show that non-experts correct the anchoring forewarning after a forewarning, whereas experts only correct with a supplementary confidence reduction. As the anchoring effect appears in many applied situations in which experts come to important decisions, this may have severe consequences. Additionally, as described above, there are reasons to believe that this may also be true for other decision biases. One option to counter this risk could be an involvement of non-experts in far reaching decisions, because they are more open to forewarnings. This idea supports positions in the recent public debate about lessons learned from the financial crisis, which actively promote more direct democracy and civic participation. Important economic and political decisions should not exclusively be made by self-assertive experts (in their field), but should also involve the participation of non-experts, who might be more open to a critical view of their own capabilities.

However, if an involvement of non-experts includes a group discussion, a critical issue needs to be addressed. Research about group discussions has repeatedly demonstrated a tendency to focus on shared information and to exclude unshared information (e.g., Gigone & Hastie, 1993; Larson, Christensen, Abbott, & Franz, 1996; Larson, Foster-Fishman, & Keys, 1994). This tendency could also be detrimental for a beneficial involvement of forewarned non-experts in decisions. They might possibly hold back their unshared viewpoint about the necessity to correct for cognitive biases in the discussion with experts. A strategy to counter the focus on shared information is the assignment of group members to specific knowledge domains at the onset of the discussion (Stasser et al., 1995; Stasser et al., 2000; Stewart & Stasser, 1995). It might therefore be crucial to inform experts and non-experts participating in the decision about the capability of non-experts to be more aware of cognitive biases, so that this strong point is appreciated and respected by all.

To ensure that the involvement of non-experts in important decisions is in fact a promising strategy to reduce the anchoring effect, additional research should clarify whether non-experts can really prevail with their more critical viewpoint when interacting with experts and whether their self-critical tendencies indeed weigh out their inferiority in other aspects.

CONCLUSION

Biases have been a very popular field of research during the past decades and are attracting increasing interest in applied contexts (e.g., Kaufmann, Michel, & Carter, 2009). In contrast to this interest and the broad practical consequences, effective and applicable debiasing methods are scarce (Lilienfeld et al., 2009).

The results of my four experiments show that for experts, forewarnings are ineffective in the case of anchoring without an additional confidence reduction. As outlined above, it is likely that the reduction of other biases by forewarnings is also hindered by heightened confidence when the recipients of these forewarnings are experts. Taken together, informing and forewarning experts about biases is obviously an often-used debiasing strategy (Wilson et al., 2002) that is not effective in the case of anchoring and may be similarly fruitless in the case of other biases.

I am not only able to demonstrate the ineffectiveness of this widespread debiasing method in the case of experts, but also to offer measures to make forewarnings effective. A reduction of confidence before giving the forewarning is an effective and feasible way to debias. Through training interventions or even books, reducing confidence in a similar way and informing people about the anchoring effect or other biases would be easy to realize. What do the reported studies reveal about the professional investors whose decisions may be biased by past stock prices (Mussweiler & Schneller, 2003)? A simple warning may fail to overcome their excessive confidence. Without reducing their overconfidence, the warnings directed at financial experts may remain unheeded, just like they did in the last financial crisis.

References

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, *108*, 441–485. doi:10.1037/0096-3445.108.4.441

Alloy, L. B., & Abramson, L. Y. (1982). Learned helplessness, depression, and the illusion of control. *Journal of Personality and Social Psychology*, *42*, 1114–1126. doi:10.1037/0022-3514.42.6.1114

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning & Verbal Behavior*, *22*, 261–295. doi:10.1016/S0022-5371(83)90201-3

Anderson, J. R. (2000). *Learning and memory: An integrated approach (2nd ed.)*. Hoboken, NJ: John Wiley & Sons Inc.

Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin*, *110*, 486–498. doi:10.1037/0033-2909.110.3.486

Arkes, H. R. (2003). Psychology in Washington: The nonuse of psychological research at two federal agencies. *Psychological Science*, *14*, 1–6. doi:10.1111/1467-9280.01410

Arkes, H. R., Christensen, C., Lai, C., & Blumer, C. (1987). Two methods of reducing overconfidence. *Organizational Behavior and Human Decision Processes*, *39*, 133–144. doi:10.1016/0749-5978(87)90049-5

Arkes, H. R., Dawes, R. M., & Christensen, C. (1986). Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes*, *37*, 93–110. doi:10.1016/0749-5978(86)90046-4

Armor, D. A. (1999). *The illusion of objectivity: A bias in the perception of freedom from* *bias.* Dissertation Abstracts International: Section B: The Sciences and Engineering, 59(9-B), 5163.

Bassarab, C. (2011). *Understanding how the leaders of three failed investment banks mismanaged early warning signals during the 2008 credit crisis*. Doctoral dissertation, Available from ProQuest Information & Learning. (UMI No. 3436965)

Blankenship, K. L., Wegener, D. T., Petty, R. E., Detweiler-Bedell, B., & Macy, C. L. (2008). Elaboration and consequences of anchored estimates: An attitudinal perspective on numerical anchoring. *Journal of Experimental Social Psychology*, *44*, 1465–1476. doi:10.1016/j.jesp.2008.07.005

Brewer, N. T., & Chapman, G. B. (2002). The fragile basic anchoring effect. *Journal of Behavioral Decision Making*, *15*, 65–77. doi:10.1002/bdm.403

Chapman, G. B., & Bornstein, B. H. (1996). The more you ask for, the more you get: Anchoring in personal injury verdicts. *Applied Cognitive Psychology*, *10*, 519–540.

Chapman, G. B., & Johnson, E. J. (1994). The limits of anchoring. *Journal of Behavioral Decision Making*, *7*, 223–242. doi:10.1002/bdm.3960070402

Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes*, *79*, 115–153. doi:10.1006/obhd.1999.2841

Chapman, G. B., & Johnson, E. J. (2002). Incorporating the irrelevant: Anchors in judgments of belief and value. In T. Gilovich, D. Griffin, & D. Kahnemann (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 120–138). New York, NY: Cambridge University Press.

Chi, M. T. H. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21–30). New York, NY: Cambridge University Press.

Clayson, D. E. (2005). Performance overconfidence: Metacognitive effects or misplaced student expectations? *Journal of Marketing Education*, *27*, 122–129. doi:10.1177/0273475304273525

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428. doi:10.1037/0033-295X.82.6.407

Craik, F. I. M. (2002). Levels of processing: Past, present . . . and future? *Memory, 10*, 305–318. doi:10.1080/09658210244000135

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, *11*, 671–684. doi:10.1016/S0022-5371(72)80001-X

Critcher, C. R., & Gilovich, T. (2008). Incidental environmental anchors. *Journal of Behavioral Decision Making, 21,* 241–251. doi:10.1002/bdm.586

Dawes, R. M. (2001). *Everyday irrationality: How pseudo-scientists, lunatics, and the rest of us systematically fail to think rationally.* Boulder, CO: Westview Press.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology*, *63*, 568–584. doi:10.1037/0022-3514.63.4.568

De Dreu, C. K. W., & Van Kleef, G. A. (2004). The influence of power on the information search, impression formation, and demands in negotiation. *Journal of Experimental*

Social Psychology, *40*, 303–319. doi:10.1016/j.jesp.2003.07.004

Ehrlinger, J., Gilovich, T., & Ross, L. (2005). Peering into the bias blind spot: People's assessments of bias in themselves and others. *Personality and Social Psychology Bulletin*, *31*, 680–692. doi:10.1177/0146167204271570

Englich, B. (2006). Blind or biased? Justitia's susceptibility to anchoring effects in the courtroom based on given numerical representations. *Law & Policy*, *28*, 497–514. doi:10.1111/j.1467-9930.2006.00236.x

Englich, B. (2008). When knowledge matters—differential effects of available knowledge in standard and basic anchoring tasks. *European Journal of Social Psychology*, *38*, 896–904. doi:10.1002/ejsp.479

Englich, B., & Mussweiler, T. (2001). Sentencing under uncertainty: Anchoring effects in the courtroom. *Journal of Applied Social Psychology*, *31*, 1535–1551. doi:10.1111/j.1559-1816.2001.tb02687.x

Englich, B., Mussweiler, T., & Strack, F. (2005). The last word in court—A hidden disadvantage for the defense. *Law and Human Behavior*, *29*, 705–722. doi:10.1007/s10979-005-8380-7

Englich, B., Mussweiler, T., & Strack, F. (2006). Playing dice with criminal sentences: The influence of irrelevant anchors on experts' judicial decision making. *Personality and Social Psychology Bulletin*, *32*, 188–200. doi:10.1177/0146167205282152

Englich, B., & Soder, K. (2009). Moody experts—How mood and expertise influence judgmental anchoring. *Judgment and Decision Making*, *4*, 41–50.

Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors. *Psychological Science*, *12*, 391–396. doi:10.1111/1467-9280.00372

Epley, N., & Gilovich, T. (2004). Are adjustments insufficient? *Personality and Social Psychology Bulletin*, *30*, 447–460. doi:10.1177/0146167203261889

Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: Differential effects of forewarning and incentives on self-generated and externally provided anchors. *Journal of Behavioral Decision Making*, *18*, 199–212. doi:10.1002/bdm.495

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science*, *17*, 311–318. doi:10.1111/j.1467-9280.2006.01704.x

Epley, N., & Gilovich, T. (2010). Anchoring unbound. *Journal of Consumer Psychology*, *20*, 20–24. doi:10.1016/j.jcps.2009.12.005

Fast, N. J., Sivanathan, N., Mayer, N. D., & Galinsky, A. D. (2012). Power and overconfident decision-making. *Organizational Behavior and Human Decision Processes*, *117,* 249–260. doi:10.1016/j.obhdp.2011.11.009

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fischer, J., Fischer, P., Englich, B., Aydin, N., & Frey, D. (2011). Empower my decisions: The effects of power gestures on confirmatory information processing. *Journal of Experimental Social Psychology*, *47,* 1146–1154. doi:10.1016/j.jesp.2011.06.008

Fischer, P., Lea, S., Kastenmüller, A., Greitemeyer, T., Fischer, J., & Frey, D. (2011). The process of selective exposure: Why confirmatory information search weakens over time. *Organizational Behavior and Human Decision Processes*, *114*, 37–48. doi:10.1016/j.obhdp.2010.09.001

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). Cambridge, England: Cambridge University Press.

Frederick, S., Kahneman, D., & Mochon, D. (2010). Elaborating a simpler theory of anchoring. *Journal of Consumer Psychology*, *20*, 17–19. doi:10.1016/j.jcps.2009.12.004

Frederick, S., & Mochon, D. (2011, July 18). A scale distortion theory of anchoring. *Journal of Experimental Psychology: General*, Advance online publication. doi:10.1037/a0024006

Galasso, A., & Simcoe, T. S. (2011). CEO overconfidence and innovation. *Management Science*, *57*, 1469–1484. doi:10.1287/mnsc.1110.1374

Galinsky, A. D., & Mussweiler, T. (2001). First offers as anchors: The role of perspective-taking and negotiator focus. *Journal of Personality and Social Psychology*, *81*, 657–669. doi:10.1037/0022-3514.81.4.657

Gervais, S., Heaton, J. B., & Odean, T. (2002). The positive role of overconfidence and optimism in investment policy. Available at SSRN: http://ssrn.com/abstract=1854465 or http://dx.doi.org/10.2139/ssrn.1854465

Gigone, D., & Hastie, R. (1993). The common knowledge effect: Information sharing and group judgment. *Journal of Personality and Social Psychology*, *65*, 959–974. doi:10.1037/0022-3514.65.5.959

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and* *semantics. Vol. 3: Speech acts* (pp. 41–58). New York, NY: Academic Press.

Guthrie, C., Rachlinski, J. J., & Wistrich, A. J. (2001). Inside the judicial mind. *Cornell Law Review*, *86*, 777–830.

Heath, C., & Tversky, A. (1990). Preference and belief: Ambiguity and competence in choice under uncertainty. In K. Borcherding, O. I. Larichev, & D. M. Messick (Eds.), *Contemporary issues in decision making.* (pp. 93–123). Oxford, England: North-Holland.

Von Helversen, B., Wilke, A., Johnson, T., Schmid, G., & Klapp, B. (2011). Performance benefits of depression: Sequential decision making in a healthy sample and a clinically depressed sample. *Journal of Abnormal Psychology*, *120*, 962–968. doi:10.1037/a0023238

Hinds, P. J. (1999). The curse of expertise: The effects of expertise and debiasing methods on predictions of novice performance. *Journal of Experimental Psychology Applied*, *5*, 205–221.

Hogarth, J. (1971). *Sentencing as a human process*. Toronto, Canada: University of Toronto Press.

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*. New Haven, CT: Yale University Press.

Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin*, *21*, 1161–1166. doi:10.1177/01461672952111004

Janiszewski, C., & Uy, D. (2008). Precision of the anchor influences the amount of adjustment. *Psychological Science*, *19*, 121–127. doi:10.1111/j.1467-

9280.2008.02057.x

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*, 697–720. doi:10.1037/0003-066X.58.9.697

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, *103*, 582–591. doi:10.1037/0033-295X.103.3.582

Kaufmann, L., Michel, A., & Carter, C. R. (2009). Debiasing strategies in supply management decision-making. *Journal of Business Logistics*, *30*, 85–106. doi:10.1002/j.2158-1592.2009.tb00100.x

Kaustia, M., Alho, E., & Puttonen, V. (2008). How much does expertise reduce behavioral biases? The case of anchoring effects in stock return estimates. *Financial Management*, *37*, 391–411.

Klayman, J., Soll, J. B., González-Vallejo, C., & Barlas, S. (1999). Overconfidence: It depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes*, *79*, 216–247. doi:10.1006/obhd.1999.2847

Knoll, M. A. Z. M. (2010). *The effects of expertise on the hindsight bias*. Doctoral dissertation, Retrieved from http://etd.ohiolink.edu/.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning & Memory*, *6*, 107–118. doi:10.1037/0278-7393.6.2.107

Krueger, J. I., & Funder, D. C. (2004). Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences*, *27*, 313–327. doi:10.1017/S0140525X04000081

Kruger, J. (1999). Lake wobegon be gone! The “below-average effect” and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology*, *77*, 221–232. doi:10.1037/0022-3514.77.2.221

Kruger, J., & Gilovich, T. (1999). “Naive cynicism” in everyday theories of responsibility assessment: On biased assumptions of bias. *Journal of Personality and Social Psychology, 76*, 743–753. doi:10.1037/0022-3514.76.5.743

Larrick, R. P. (2004). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making.* (pp. 316–337). Malden, MA: Blackwell Publishing.

Larrick, R. P., Burson, K. A., & Soll, J. B. (2007). Social comparison and confidence: When thinking you’re better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes, 102*, 76–94.

Larson, J. R., Christensen, C., Franz, T. M., & Abbott, A. S. (1998). Diagnosing groups: The pooling, management, and impact of shared and unshared case information in team-based medical decision making. *Journal of Personality and Social Psychology, 75*, 93–108. doi:10.1037/0022-3514.75.1.93

Larson, J. R., Foster-Fishman, P. G., & Keys, C. B. (1994). Discussion of shared and unshared information in decision-making groups. *Journal of Personality and Social Psychology*, *67*, 446–461. doi:10.1037/0022-3514.67.3.446

LeBoeuf, R. A., & Shafir, E. (2009). Anchoring on the “here” and “now” in time and distance judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 81–93. doi:10.1037/a0013665

Lichtenstein, S., & Fischhoff, B. (1980). Training for calibration. *Organizational Behavior*

& Human Performance, 26, 149–171. doi:10.1016/0030-5073(80)90052-5

Lilienfeld, S. O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on Psychological Science*, *4,* 390–398. doi:10.1111/j.1745-6924.2009.01144.x

Lockhart, R. S. (2002). Levels of processing, transfer-appropriate processing, and the concept of robust encoding. *Memory*, *10*, 397–403. doi:10.1080/09658210244000225

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*, 2098–2109. doi:10.1037/0022-3514.37.11.2098

Ludwig, S., Wichardt, P. C., & Wickhorst, H. (2011, May 27). On the positive effects of overconfident self-perception in teams. (May 27, 2011). Available at SSRN: http://ssrn.com/abstract=1854465 or http://dx.doi.org/10.2139/ssrn.1854465.

Lundgren, S. R., & Prislin, R. (1998). Motivated cognitive processing and attitude change. *Personality and Social Psychology Bulletin*, *24*, 715–726. doi:10.1177/0146167298247004

McKenzie, C., Liersch, M., & Yaniv, I. (2008). Overconfidence in interval estimates: What does expertise buy you? *Organizational Behavior and Human Decision Processes*, *107,* 179–191. doi:10.1016/j.obhdp.2008.02.007

Montier, J. (2007). *Behavioural investing: A practitioners guide to applying behavioural finance*. Chichester, UK: John Wiley & Sons.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, *115*, 502–517. doi:10.1037/0033-295X.115.2.502

Moore, D. A., Kurtzberg, T. R., Fox, C. R., & Bazerman, M. H. (1999). Positive illusions and forecasting errors in mutual fund investment decisions. *Organizational Behavior and Human Decision Processes, 79*, 95–114.

Moore, D. A., & Swift, S. A. (2011). The three faces of overconfidence in organizations. In D. De Cremer, R. van Dick, & J. K. Murnighan (Eds.), *Social Psychology and Organizations* (pp. 147–184). New York, NY: Routledge/Taylor & Francis Group.

Mussweiler, T. (2001). The durability of anchoring effects. *European Journal of Social Psychology, 31*, 431–442. doi:10.1002/ejsp.52

Mussweiler, T., & Englich, B. (2005). Subliminal anchoring: Judgmental consequences and underlying mechanisms. *Organizational Behavior and Human Decision Processes*, *98,* 133–143. doi:10.1016/j.obhdp.2004.12.002

Mussweiler, T., Englich, B., & Strack, F. (2004). Anchoring effect. In R. Pohl (Ed.), *Cognitive illusions—A handbook on fallacies and biases in thinking, judgment, and memory* (pp. 183–200). London, UK: Psychology Press.

Mussweiler, T., & Schneller, K. (2003). "What goes up must come down"—how charts influence decisions to buy and sell stocks. *Journal of Behavioral Finance, 4*, 121–130. doi:10.1207/S15427579JPFM0403

Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology, 35,* 136–164. doi:10.1006/jesp.1998.1364

Mussweiler, T., & Strack, F. (2000a). Numeric judgments under uncertainty: The role of knowledge in anchoring. *Journal of Experimental Social Psychology*, *36*, 495–518. doi:10.1006/jesp.1999.1414

Mussweiler, T., & Strack, F. (2000b). The use of category and exemplar knowledge in the solution of anchoring tasks. *Journal of Personality and Social Psychology*, *78*, 1038–1052. doi:10.1037/0022-3514.78.6.1038

Mussweiler, T., & Strack, F. (2001a). Considering the impossible: Explaining the effects of implausible anchors. *Social Cognition*, *19*, 145–160. doi:10.1521/soco.19.2.145.20705

Mussweiler, T., & Strack, F. (2001b). The semantics of anchoring. *Organizational Behavior and Human Decision Processes*, *86*, 234–255. doi:10.1006/obhd.2001.2954

Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin*, *26*, 1142–1150. doi:10.1177/01461672002611010

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. doi:10.1016/0749-5978(87)90046-X

Pallier, G., Wilkinson, R., Danthir, V., Kleitman, S., Knezevic, G., Stankov, L., & Roberts, R. D. (2002). The role of individual differences in the accuracy of confidence judgments. *Journal of General Psychology*, *129*, 257–299. doi:10.1080/00221300209602099

Partridge, A., & Eldridge, W. B. (1974). *The second circuit sentencing study*. Washington, DC: The Federal Judicial Center.

Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology, 46*, 69–81. doi:10.1037/0022-3514.46.1.69

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. doi:10.1006/jesp.1993.1007

Phillips, L. D. (1973). *Bayesian statistics for social scientists*. London, UK: Thomas Nelson.

Piper, N. (2009). Der Prophet des Untergangs. *sueddeutsche.de*, Retrieved from http://www.sueddeutsche.de/wirtschaft/oekonom–nouriel–roubini–der–prophet–des–untergangs–1.488149–.

Plous, S. (1993). *The psychology of judgment and decision making*. McGraw-Hill series in social psychology. New York, NY: Mcgraw-Hill Book Company.

Plous, S. (1995). A comparison of strategies for reducing interval overconfidence in group judgments. *Journal of Applied Psychology, 80*, 443–454. doi:10.1037/0021-9010.80.4.443

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879–891. doi:10.3758/BRM.40.3.879

Pronin, E. (2008). How we see ourselves and how we see others. *Science*, *320*, 1177–1180. doi:10.1126/science.1154199

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*, *111*, 781–799. doi:10.1037/0033-295X.111.3.781

Pronin, E., & Kugler, M. B. (2007). Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. *Journal of Experimental Social Psychology*, *43*, 565–578.

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, *28*, 369–381. doi:10.1177/0146167202286008

Rajan, R. G. (2005). Has financial development made the world riskier? *National Bureau of Economic Research Working Paper Series*, *No. 11728*. http://www.nber.org/papers/w11728

Reitsma-van Rooijen, M., & Daamen, D. D. L. (2006). Subliminal anchoring: The effects of subliminally presented numbers on probability estimates. *Journal of Experimental Social Psychology*, *42*, 380–387. doi:10.1016/j.jesp.2005.05.001

Ritov, I. (1996). Anchoring in simulated competitive market negotiation. *Organizational Behavior and Human Decision Processes*, *67*, 16–25. doi:10.1006/obhd.1996.0062

Roubini, N. (2006, September 7). EconoMonitor Flashback: Roubini’s IMF Speech. *EconoMonitor*, Retrieved from: http://www.economonitor.com/nouriel/2010/09/02/economonitor–flashback–roubinis–imf–speech–september–7–2006/.

Russo, J. E. (2010). Understanding the effect of a numerical anchor. *Journal of Consumer Psychology*, *20*, 25–27. doi:10.1016/j.jcps.2009.12.006

Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation. In M. P. Zanna (Ed.), *Advances in experimental social psychology, Vol. 26.* (pp. 123–162). San Diego, CA: Academic Press.

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *47A*, 631–650.

Sharp, G. L., Cutler, B. L., & Penrod, S. D. (1988). Performance feedback improves the resolution of confidence judgments. *Organizational Behavior and Human Decision Processes*, *42*, 271–283. doi:10.1016/0749-5978(88)90001-5

Shiller, R. J. (2006). *Irrational exuberance* (2nd ed.). Princeton, NJ: Princeton University Press.

Sieck, W. R., & Arkes, H. R. (2005). The recalcitrance of overconfidence and its contribution to decision aid neglect. *Journal of Behavioral Decision Making*, *18*, 29–53. doi:10.1002/bdm.486

Simmons, J. P., LeBoeuf, R. A., & Nelson, L. D. (2010). The effect of accuracy motivation on anchoring and adjustment: Do people adjust from provided anchors? *Journal of Personality and Social Psychology*, *99*, 917. doi:10.1037/a0021540

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *69*, 99–118. doi:10.2307/1884852

Simon, H. A. (1957). *Models of man; social and rational.* Oxford, UK: Wiley.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1987). Behavioral decision theory perspectives on protective behavior. In N. D. Weinstein (Ed.), *Taking care: Understanding and encouraging self-protective behavior* (pp. 14–41). New York, NY: Cambridge University Press.

Smith, A. R. (2011). *Exploring the relationship between knowledge and anchoring effects: Is the type of knowledge important?* Doctoral dissertation, University of Iowa. http://ir.uiowa.edu/etd/1264

Soll, J. B., & Klayman, J. (2004). Overconfidence in interval estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 299–314. doi:10.1037/0278-7393.30.2.299

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*, 645–726. doi:10.1017/S0140525X00003435

Stasser, G., Stewart, D. D., & Wittenbaum, G. M. (1995). Expert roles and information exchange during discussion: The importance of knowing who knows what. *Journal of Experimental Social Psychology*, *31*, 244–265. doi:10.1006/jesp.1995.1012

Stasser, G., Vaughan, S. I., & Stweart, D. D. (2000). Pooling unshared information: The benefits of knowing how access to information is distributed among group members. *Organizational Behavior and Human Decision Processes*, The psychological foundations of knowledge transfer in organizations, *82*, 102–116. doi:10.1006/obhd.2000.2890

Stewart, D. D., & Stasser, G. (1995). Expert role assignment and information sampling during collective recall and decision making. *Journal of Personality and Social Psychology*, *69*, 619–628. doi:10.1037/0022-3514.69.4.619

Stewart, N. (2009). The cost of anchoring on credit-card minimum repayments. *Psychological Science*, *20*, 39–41. doi:10.1111/j.1467-9280.2008.02255.x

Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF) [Multidimensional comfort questionnaire]*. Göttingen, Germany: Hogrefe.

Strack, F., & Hannover, B. (1996). Awareness of influence as a precondition for implementing correctional goals. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 579–596). New York, NY: Guilford Press.

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, *73*, 437–446. doi:10.1037/0022-3514.73.3.437

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, *47*, 143–148.

Taleb, N. N. (2008). *The black swan: The impact of the highly improbable*. London, UK: Penguin UK.

Törngren, G., & Montgomery, H. (2004). Worse than chance? Performance and confidence among professionals and laypeople in the stock market. *Journal of Behavioral Finance*, *5*, 148–153. doi:10.1207/s15427579jpfm0503_3

Trafimow, D., & Sniezek, J. A. (1994). Perceived expertise and its effect on confidence. *Organizational Behavior and Human Decision Processes*, *57*, 290–302.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. doi:10.1126/science.185.4157.1124

Vázquez, C. (1987). Judgment of contingency: Cognitive biases in depressed and nondepressed subjects. *Journal of Personality and Social Psychology, 52*, 419–431.

doi:10.1037/0022-3514.52.2.419

Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology*, *68*, 36–51. doi:10.1037/0022-3514.68.1.36

Wegener, D. T., Petty, R. E., Blankenship, K. L., & Detweiler-Bedell, B. (2010a). Elaboration and numerical anchoring: Breadth, depth, and the role of (non-)thoughtful processes in anchoring theories. *Journal of Consumer Psychology*, *20*, 28–32. doi:10.1016/j.jcps.2009.12.007

Wegener, D. T., Petty, R. E., Blankenship, K. L., & Detweiler-Bedell, B. (2010b). Elaboration and numerical anchoring: Implications of attitude theories for consumer judgment and decision making. *Journal of Consumer Psychology*, *20*, 5–16. doi:10.1016/j.jcps.2009.12.003

Wegener, D. T., Petty, R. E., Detweiler-Bedell, B. T., & Jarvis, W. B. G. (2001). Implications of attitude change theories for numerical anchoring: Anchor plausibility and the limits of anchor effectiveness. *Journal of Experimental Social Psychology*, *37*, 62–69. doi:10.1006/jesp.2000.1431

Wetzel, C. G., Wilson, T. D., & Kort, J. (1981). The halo effect revisited: Forewarned is not forearmed. *Journal of Experimental Social Psychology*, *17*, 427–439. doi:10.1016/0022-1031(81)90049-4

Whyte, G., & Sebenius, J. K. (1997). The effect of multiple anchors on anchoring in individual and group judgment. *Organizational Behavior and Human Decision Processes*, *69*, 75–85. doi:10.1006/obhd.1996.2674

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction:

Unwanted influences on judgments and evaluations. *Psychological Bulletin, 116,* 117–142. doi:10.1037/0033-2909.116.1.117

Wilson, T. D., Centerbar, D. B., & Brekke, N. (2002). Mental contamination and the debiasing problem. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 185–200). New York, NY: Cambridge University Press.

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. doi:10.1037/0096-3445.125.4.387

Wong, K. F. E., & Kwong, J. Y. Y. (2000). Is 7300 m equal to 7.3 km? Same semantics but different anchoring effects. *Organizational Behavior and Human Decision Processes*, *82,* 314–333. doi:10.1006/obhd.2000.2900

Wright, W. F., & Anderson, U. (1989). Effects of situation familiarity and financial incentives on use of the anchoring and adjustment heuristic for probability assessment. *Organizational Behavior and Human Decision Processes*, *44,* 68–82. doi:10.1016/0749-5978(89)90035-6