Cognitive Psychology 90 (2016) 1–47

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

The anchor integration model: A descriptive model of anchoring effects ☆

Brandon M. Turner [a,*], Dan R. Schley [b]

[a] Department of Psychology, The Ohio State University, United States
[b] Rotterdam School of Management, Erasmus University, Netherlands

ARTICLE INFO

Article history:
Accepted 19 July 2016

Keywords:
Anchoring
Judgments
Modeling
Estimation

ABSTRACT

Few experimental effects in the psychology of judgment and decision making have been studied as meticulously as the anchoring effect. Although the existing literature provides considerable insight into the psychological processes underlying anchoring effects, extant theories up to this point have only generated qualitative predictions. While these theories have been productive in advancing our understanding of the underlying anchoring process, they leave much to be desired in the interpretation of specific anchoring effects. In this article, we introduce the Anchor Integration Model (AIM) as a descriptive tool for the measurement and quantification of anchoring effects. We develop two versions the model: one suitable for assessing between-participant anchoring effects, and another for assessing individual differences in anchoring effects. We then fit each model to data from two experiments, and demonstrate the model's utility in describing anchoring effects.

1. Introduction

As technology and innovation have brought a new age of intellectual prosperity, we find ourselves with a multitude of quantitative information. For years, psychologists have tried to understand how

☆ This research was supported by a National Science Foundation Graduate Research Fellowship to the second author, and a National Science Foundation Research Grant SES-1047757 to Ellen Peters. We thank Tim Pleskac, Ellen Peters, Michael DeKay, Gretchen Chapman, Duane Wegener, and Carly Muller for helpful comments on earlier versions of this manuscript.

* Corresponding author.
E-mail address: turner.826@gmail.com (B.M. Turner).

http://dx.doi.org/10.1016/j.cogpsych.2016.07.003
0010-0285/

incidental numbers influence judgments. Anchoring effects were brought to the forefront of psychological research by Kahneman and Tversky (1974), although earlier examples of a similar phenomenon were demonstrated in previous research on preference reversals (Lichtenstein & Slovic, 1971; Slovic, 1967; Slovic & Lichtenstein, 1968). In their original article, Kahneman and Tversky first asked participants to indicate whether the percentage of African nations in the United Nations was greater than or less than an arbitrary number, derived from spinning a wheel of fortune. This arbitrary number, the anchor, was set to be either high (e.g., 65%) or low (e.g., 10%). Participants were then asked to give their best estimate of the percentage of African nations in the United Nations. Results indicated that participants' estimates assimilated to the provided anchor value, such that the mean estimate of participants in the high-anchor condition was 45% and the mean estimate of participants in the low-anchor condition was 25%.

Over the past forty years, anchoring research has received extensive attention because of its robust effect sizes and its broad applicability in a myriad of domains. One such example was provided by Northcraft and Neale (1987), in which students and real estate appraisers toured a home and appraised its value. Although each participant had relevant experiential information about the home, both groups showed a significant correlation between their appraisal and the anchor value. In another example, Plous (1989) demonstrated that students' estimates of the likelihood of nuclear war were significantly influenced by arbitrary anchors. Other common paradigms include the pricing of gambles (Carlson, 1990; Chapman & Johnson, 1994, 1999; Johnson & Schkade, 1989), self efficacy (Cervone & Peake, 1986), negotiations (Galinsky & Mussweiler, 2001), judicial verdicts (Chapman & Bornstein, 1996; Englich & Mussweiler, 2001; Englich, Mussweiler, & Strack, 2006), consumer decisions and willingness-to-pay (Ariely, Loewenstein, & Prelec, 2003; Green, Jacowitz, Kahneman, & McFadden, 1998; Simonson & Drolet, 2004; Stewart, 2009; Wansink, Kent, & Hoch, 1998), debt repayment (Navarro-Martinez et al., 2011; Stewart, 2009), and general knowledge questions (Jacowitz & Kahneman, 1995; McElroy & Dowd, 2007; Strack & Mussweiler, 1997).

1.1. The problem with the anchoring effect

Although it is clear that anchoring effects have widespread implications, less is known about the cognitive mechanisms that drive these effects. Over the past few decades, theories of the anchoring-effect mechanism have risen, but unfortunately, extant theories have not fallen. To be clear, having several non-mutually exclusive theories is acceptable, if not favorable, when there are multiple unique cognitive processes involved in the decision and it is clear under which conditions each theoretical mechanism plays a more versus less critical role in the decision process. We suggest that the anchoring literature lacks on this latter point, leaving researchers and practitioners with very little insight about which extant theory should apply to a given context, thereby limiting anchoring-effect predictions. Currently, there are five major theories – reviewed below – some with additional minor offshoots, that provide insight about how and why individuals assimilate judgments to presented anchor values. These theories provide psychological mechanisms that explain the processes underlying the anchoring effect and potentially relevant boundary conditions for the effect (for a review, see Furnham & Boo, 2011). Nevertheless, after 4 decades of research, what *exactly* do we know about the anchoring effect and how to predict it?

Anchoring effects often involve the presentation of a quantitative anchor, from which participants produce a quantitative response (although there are exceptions, e.g., Oppenheimer, LeBoeuf, & Brewer, 2008). Because the anchoring effect refers to the assimilation of a quantitative judgment to a quantitative anchor, qualitative theories are incapable of fully articulating the effect. In particular, extant theories provide considerable insight into the psychological processes that ostensibly produce the assimilation of individuals' judgments toward a presented anchor value, though they offer little insight into the magnitude of said assimilation. For example, consider a question in which participants are asked to estimate the length of the Mississippi river. Current theories provide qualitative predictions about how the presence of a high or low anchor will affect the shift of the judgments toward the anchor value, but they do not describe the distribution of these judgments. Hence, current theories of anchoring effects are incapable of making quantitative predictions for say, an anchor value of 3500 miles versus an anchor value of 1500 miles.

The extensive anchoring literature provides considerable evidence for the robustness of the anchoring effect across substantive domains. If the anchoring effect can reliably influence important decisions like home prices (Northcraft & Neale, 1987), judicial decisions (Chapman & Bornstein, 1996; Englich & Mussweiler, 2001; Englich et al., 2006), and debt repayment (Navarro-Martinez et al., 2011; Stewart, 2009), then the ability to make quantitative predictions (i.e., the amount of assimilation for any given anchor) about anchoring effects would be extremely useful. In this article, we attempt to provide a descriptive modeling tool for the study of anchoring effects.

Considerable attention has recently focused on the contribution of behavioral research as a tool to allow policy makers to "nudge" individuals' judgments and decisions in a manner that facilitates socially desirable outcomes (Allcott & Mullainathan, 2010; Johnson & Goldstein, 2003; Thaler & Sunstein, 2008). Take for instance the proportion of income that American consumers save. The savings rate has large macroeconomic consequences, with deleterious outcomes associated with both savings rates that are too low (i.e., the aggregate costs of interest decreases consumers' subsequent buying power, and as a consequence, gross domestic product) and too high (i.e., lower rates of consumer consumption generally decrease gross domestic product). Here, policy recommendations intended to nudge consumer spending would act as anchors (e.g., telling individuals that the ideal savings rate is 7% for 25–40 year-olds). Extant anchoring research offers intuition about what types of anchors produce relatively larger and smaller anchoring effects, but offer little insight into which anchors might produce the most optimal outcomes. Because there are more and less optimal saving rates, tools that quantify important behavioral effects offer policy makers the ability to potentially introduce optimal nudges (Urminsky & Goswami, 2015).

In subsequent sections we will begin by reviewing extant theories as a means of guiding the development of our model. Next, we present two versions of the Anchor Integration Model (AIM) for the prediction of quantitative judgments. We will first present a version of AIM developed to model anchoring effects on a between-participant level (e.g., providing a group of participants a particular anchor value and estimating the distribution of corresponding judgments). Following, we will present a within-participant version of AIM meant to characterize how an individual anchor influences an individual's judgment. As a first endeavor into quantitatively evaluating anchoring effects, the goal of the present article is to develop a descriptive model of the data, rather than a mechanistic model of the underlying psychological process.

2. Current theories of anchoring and a general framework for AIM

Although AIM is not intended to be a process model, reviewing current theories of the psychological mechanism at play will be useful in guiding the development of our descriptive model.

2.1. Anchoring-and-adjustment theory

The first theory in this line of research – Anchoring-and-adjustment theory – assumes that participants anchor their judgments on the provided anchor value and insufficiently adjust their estimates away from the anchor (Kahneman & Tversky, 1974). Consider the length of the Mississippi River. According to Anchoring-and-adjustment, individuals will begin at a given anchor value (e.g., 3500 miles) and make a series of incremental adjustments, pausing after each adjustment to decide whether or not to continue adjusting. The anchoring effect is thought to occur because individuals insufficiently adjust away from the anchors, resulting in estimates that are skewed toward the anchor value (Epley & Gilovich, 2006).

Within the literature, two primary accounts are used to explain this insufficient adjustment. The first account argues that individuals have a plausible set of values for all judgments and that estimates are derived from the first value within their plausible interval (Epley & Gilovich, 2006; Quattrone, 1982; Quattrone, Lawrence, Finkel, & Andrus, 1984). The second account of Anchoring-and-adjustment focuses on the attention-demanding nature of adjustment. According to this account, adjustment is an elaborative process in which an individual must consider each adjustment and choose whether to continue elaborating (i.e., adjusting to a new value) or settle on the current value

(Epley, 2004; Epley & Gilovich, 2006; Gilbert, 2002). Because attention is a limited resource, insufficient adjustment is a product of the cumulative costs of adjusting.

2.2. Numeric and magnitude priming

Initially set forth by Wilson, Houston, Etling, and Brekke (1996), the Numeric-priming account suggests that *any* number present in the judgment environment can act as an anchor, independent of context or relevance. For example, participants in one study expected that the price of a meal would be higher at a restaurant named "Studio 97" than at one named "Studio 17" (Critcher & Gilovich, 2008). In addition, Oppenheimer et al. (2008) demonstrated that simply drawing a long line before producing an estimate resulted in reliably larger estimates than those produced after drawing a short line. That is, drawing a line produces a sense of "bigness" or "smallness" which primes magnitude information in a subsequent judgment. The primary proposition of this account is that any magnitude information available during the focal judgment is conflated with the judgment-relevant information. For example, if an individual sees the number 3500 before being asked to judge the length of the Mississippi River, the "3500" will be added to judgment-relevant information, regardless of its actual informativeness.

2.3. The selective accessibility model

The Selective-accessibility model postulates that anchoring effects are mediated by a selective increase in the accessibility of anchor-consistent knowledge about the target item (Chapman & Johnson, 1994, 1999, 2002; Mussweiler & Strack, 1999; Mussweiler, Strack, & Pfeiffer, 2000; Strack & Mussweiler, 1997). Specifically, Selective Accessibility supposes that when making the comparative judgment in an anchoring task (e.g., is the Mississippi River greater than or less than 3500 miles?) individuals test the hypothesis that the anchor is equal to the true value of the target item. To perform this confirmatory hypothesis test, individuals *selectively* retrieve information that is consistent with the Mississippi River being 3500 miles long (see Klayman et al., 1991; Wason, 1960). When the estimate is subsequently elicited, this retrieved information is more *accessible*, biasing estimates toward the anchor value. In support of this model, Mussweiler and Strack (1999, 2000, 2001) conducted a series of studies measuring the increase in available information around the anchor value. In one study, participants who had indicated whether the average price for a German automobile was greater than or less than the high anchor of 40,000 German Marks, were faster at identifying words associated with expensive cars (e.g., Mercedes-Benz) than words associated with cheaper cars (e.g., Volkswagen) in a lexical decision task (Mussweiler & Strack, 2000). Thus, information around the anchor value is activated in memory in a manner consistent with Selective Accessibility.

2.4. The attitudinal perspective of anchoring

A recent perspective of numeric anchoring has arisen from the attitudes and persuasion literature. The Attitudinal Perspective contends that individuals process numeric anchors and persuasive messages in similar ways (Wegener, Petty, Blankenship, & Detweiler-Bedell, 2010). In particular, the Attitudinal Perspective argues that different types of anchoring can occur between thoughtful and non-thoughtful processes, in a manner consistent with the elaboration likelihood model (Petty & Cacioppo, 1986). An efficient summary was provided by Wegener et al. (2010):

> "Effortful use of anchor-consistent background knowledge provides a reasonable account for high-elaboration anchoring, but there may be a number of relatively non-thoughtful processes that contribute to low-elaboration anchoring. Numeric (magnitude) priming is one potential process, and it also seems plausible that people take at least some anchors as 'hints' as to the correct (direction of) target judgment." (Wegener et al., 2010, p. 11).

Simply put, when individuals are elaborating on the judgment, they process anchors in a manner consistent with Selective Accessibility, but when they are not elaborating, they process anchors in a manner consistent with the Numeric-priming account. One additional advantage of the Attitudinal Perspective is that it allows for anchors to serve as "hints" (Schwarz, 1994; Wegener et al., 2010; Wegener, Petty, Detweiler-Bedell, & Jarvis, 2001). In particular, participants may assume that the presented anchor provides implicit information regarding the judgment (Grice, 1975; Jacowitz & Kahneman, 1995; Schwarz, 1994); a feature not available in prior accounts. Consistent with this notion, Zhang and Schwarz (2013) found that anchors with precise numbers (e.g., 3478) resulted in larger anchoring effects than comparable round anchors (e.g., 3500) because individuals use the precision of the presented anchor as a means of assessing the quality of the anchor information. As noted by Wegener et al. (2010), an anchor from a credible source should result in a larger anchoring effect than an anchor from a less credible source.

2.5. The scale distortion theory of anchoring

The most recent addition to the growing collection of anchoring theories is the Scale-distortion theory of anchoring proposed by Frederick and Mochon (2012). According to this account, individuals map judgments to an underlying response scale. The presentation of an anchor does not shift individuals' subjective representation of the target item (Mussweiler, 2003), but informs the individual about the scale used to make the judgment. For example, when presented a low anchor for the length of the Mississippi River (e.g., 70 miles), individuals typically produce lower estimates for the river's length (e.g., 1000 miles), compared to a control condition without an anchor (e.g., 2000 miles). Scale Distortion suggests that individuals do not perceive the Mississippi River as any shorter, rather the scale used to make the judgment distorts around the anchor value (Frederick & Mochon, 2012; Mochon & Frederick, 2013). When judging the river's length, given the presence of a 70 mile anchor, 1000 miles seems sufficient to convey that the Mississippi River is very long, but without the 70 mile anchor to serve as a cue, 2000 miles seems necessary to convey the "largeness" of the length of the river. For example, Frederick and Mochon (2012) presented participants with a list of 15 animals ordered by weight. Participants were tasked with choosing an animal whose average adult weight was closest to 1000 pounds. Before doing so, half of the participants estimated the average weight of an adult wolf, the other half did not. Results indicated that participants who first judged the weight of a wolf selected examples of a 1000-pounds animal that were substantially larger than did those who did not first judge the weight of the wolf. According to the authors, because wolves are much less than 1000 pounds, if anchors (i.e., judging the weight of a wolf) distort the scale of animal weights around the anchor, then a much larger animal would be chosen as an exemplar of a 1000-pounds animal. Simply put, animals smaller than the anchor would assimilate to values higher along the scale and animals larger than the anchor would assimilate to values lower along the scale.

2.6. Integrating across theories

To the reader, the existence of five non-mutually-exclusive theories may cause concern about the efficacy of these theories or the anchoring literature more generally. We suggest that the persistence of this ever increasing number of theories is due to a misnomer within this literature. This assimilation of judgments to presented anchor values is commonly referred to as "the anchoring effect," despite a range of judgment domains and types of anchors. We believe that it would be helpful to conceptualize this not as "an effect" but as a class of effects with a common antecedent (i.e., the presence of an anchor) and a common consequence (i.e., the assimilation of judgments toward the anchor value). To facilitate clarity, we would suggest referring to "anchoring effects" as opposed to "the anchoring effect." If the body of experiments in this literature represent a class of effects, and the existing theories provide insight into the domain-general processes across paradigms, then our model may be informed by considering commonalities across the range of theories and anchoring effects.

The theories reviewed above differ considerably with regard to how the anchor influences judgments: According to Anchoring-and-adjustment, the anchor provides a meaningful starting point to begin the elaborative iteration process.[1] Numeric Priming suggest that the anchor acts as an uncontextualized information node that is assimilated into the judgment. Selective Accessibility suggests that the anchor acts as a cue that elicits a confirmatory hypothesis test; increasing the accessibility of anchor-consistent knowledge. The Attitudinal Perspective asserts that the anchor is treated as a persuasive message, such that the individual treats the anchor as a conversational hint. Lastly, the Scale-distortion theory of anchoring proposes that the anchor informs the scale with which the judgment is made. Despite these differences, the above theories share the proposition that the anchor is a source of information, in one way or another. If it is fair to assume that the theories agree that the anchor is an informational source that influences judgments, then how might we use this insight to develop our model?

We begin by attempting to identify the subcomponent processes involved in anchoring effects. Consider the most common anchoring paradigm, where participants are presented with an anchor value and are asked to make a comparative judgment (e.g., is the length of the Mississippi River less than or greater than 3500 miles?) before making the target judgment. The first component of the anchoring process is prior knowledge. For example, research indicates that individuals with more prior knowledge about the target item are less influenced by anchors, exhibiting smaller anchoring effects (e.g., Smith, Windschitl, & Bruchmann, 2013; Smith & Windschit, 2015). The second component of the anchoring process is the information provided by the anchor (see previous paragraph). The last component we consider is how the mind integrates the first and second components. If individuals have both prior information and information provided by the anchor, then how are these two sources of information integrated into a judgment? We decompose these components of the anchoring process because we believe that they may have unique psychological antecedents. We believe that identifying and measuring these subcomponents of anchoring effects is useful for understanding the contributions of prior literature and the current article. Most of the anchoring literature, particularly those which posit individual anchoring theories, tend to focus on the second component. As discussed above, most of these theories agree that the anchor provides some form of information, but they differ in terms *how* the anchor provides information. Despite the size of the anchoring literature, far less research has been dedicated to the first and third components. The current article focuses primarily on the third component. We will not propose an explanation of how the mind organizes prior information (component 1), derives information from the anchor (component 2), or integrates these two sources of information (component 3). Instead, we offer a descriptive tool for measuring the result of these latent psychological components of anchoring.

AIM is not the first model intended to capture anchoring effects. Past research on belief updating in order effects (Hogarth & Einhorn, 1992) and preference reversals (Goldstein & Einhorn, 1987; Johnson & Busemeyer, 2005) have used an anchoring-and-adjustment mechanism as a means of describing a particular effect (as an alternative, see Birnbaum & Zimmermann, 1998); though few models have focused primarily on anchoring (e.g., Bhatia & Chaudhry, 2013). Other accounts have used Bayesian mechanisms for updating, but assumed approximate inference algorithms to show basic anchoring effects (Lieder, Griffiths, & Goodman, 2012). To varying extents, these models relate in that they treat anchoring as a process of belief updating and information integration. Complementing the intuition of these models, AIM develops a more comprehensive structure for quantifying this updating process.

This notion of information updating (i.e., the third component of our anchoring framework) is neither new nor is it exclusive to anchoring and related effects. Within other areas of cognition, researchers have long studied how the mind integrates the presentation of a novel stimulus with their existing knowledge (for a review, see Anderson, 1990). If anchoring effects involve the integration of a novel stimulus with individuals' prior information, in a manner similar to other areas of cognition, then intuition from these other areas may be particularly useful in developing AIM. Within cognitive psychology and related fields, researchers not only attempt to develop theories for how the mind processes information, but also develop quantitative models that (1) allow for the prediction of effects, and

[1] We acknowledge that Anchoring-and-adjustment theory is theoretically distinct from later theories. In Anchoring-and-adjustment, the anchor is assumed to be uninformative, and only serves as a starting point for this adjustment. While the anchor is uninformative in terms of its diagnostic value (e.g., the anchor value coming from a spinning wheel of fortune), the anchor *is* informative as a cue for the beginning of the supposed adjustment process. As such, although we acknowledge functional differences between Anchoring-and-adjustment and subsequent theories, we believe that within Anchoring-and-adjustment, the anchor has informational value for the target judgment (i.e., informing individuals from *where* to start their adjustment process).

(2) provide a framework for testing theoretical predictions regarding psychological processes; we will return to the latter point in the General Discussion.

There are also many dynamic models of perception that have demonstrated the usefulness of assuming the integration of novel stimuli into existing knowledge representations. How these models regulate the integration of information provides an important distinction between them. A special class of models assume that the integration occurs within the Bayesian framework and are often referred to as "Bayesian models of cognition" (Griffiths, Kemp, & Tenenbaum, 2008). Bayesian models of cognition incorporate both prior information and the information provided by the stimuli to generate a response. These models have become very popular because they provide a convenient way to not only model the cognitive mechanisms underlying the behavior of interest, but they also provide a method for rule-based generalization of cognitive processes. Bayesian models have been used to explain lower-level cognitive processes such as visual scene perception (Yuille & Kersten, 2006), semantic memory (Steyvers, Griffiths, & Dennis, 2006), and language acquisition (Chater & Manning, 2006). For decades, research in judgment and decision making has, in large, assumed more elaborative cognitive processes (e.g., judgment and choice) to be simplifying heuristics rather than an algorithmic procedure (for an exhaustive review of this literature, see Gilovich, Griffin, & Kahneman, 2002). In contrast to these simplifying heuristics, recent research has focused on developing Bayesian models to account for more elaborative judgments such as human inductive learning (Tenenbaum, Griffiths, & Kemp, 2006), social cognition (Baker, Tenenbaum, & Saxe, 2007), and future predictions (Griffiths & Tenenbaum, 2011). Lieder et al. (2012) have even proposed a Bayesian model that uses an approximate inference strategy as a rational account of simple anchoring effects. Their model relies on the observation that humans seem to sample limited evidence when forming their representations, and in our efforts to reduce computational cost, a bias toward information cues is induced.

Connecting the discussion above and reiterating our basic framework for the subcomponents of anchoring, we propose that anchoring effects involve two primary informational inputs: (1) individuals' prior representations about a target judgment (i.e., the information the individual would have available in the absence of the anchor value), and (2) the information provided by the anchor. These two sources of information are likely integrated in a manner analogous to the processes proposed in the information-updating literature reviewed above. In the versions of AIM presented subsequently, we do not attempt to propose the underlying processes involved in each of the three components of anchoring, we only provide a descriptive tool. More directly, we remain agnostic as to *how* the anchor provides information; how the anchor provides information is discussed extensively within the existing anchoring theories. We generally assume that the integration of the anchor information with one's prior information will occur in a manner analogous to Bayesian updating. Although Bayesian updating models are one way of conceptualizing the integration, there also exists relevant alternate models with simpler forms of information integration (e.g., Turner, Van Zandt, & Brown, 2011).

The basic intuition of AIM is illustrated in Fig. 1. AIM assumes that individuals have some prior representation of any target judgment. For instance, consider an individual who believes the river to be roughly between 1000 and 3000 miles long, with their best guess at around 2000 miles. AIM conceptualizes this representation as the gray information density function in Fig. 1A. For the sake of this example we use simple Gaussian distributions, but we discuss distributional assumptions for AIM when we present the within-participant version of AIM in the subsequent section. Because individuals can have different amounts of information, or confidence, in their prior representation, differences are captured in terms of how diffuse the representations are. For example, an individual with more information about the Mississippi River being 2000 miles, or having more confidence in their information, would have a less diffuse prior representation (Fig. 1B). Conversely, individuals with less information or confidence would have a more diffuse representation (Fig. 1C).

Now consider that the individual is presented an anchor about the Mississippi River being 3500 miles long. The green[2] distributions in Fig. 1 illustrate the anchor information. For the sake of this illustration, consider the anchor information as an information density function centered on 3500 miles. This function is diffuse to illustrate the weakness of the anchor information. For instance, an anchor can pro-

[2] For interpretation of color in Fig. 1, the reader is referred to the web version of this article.

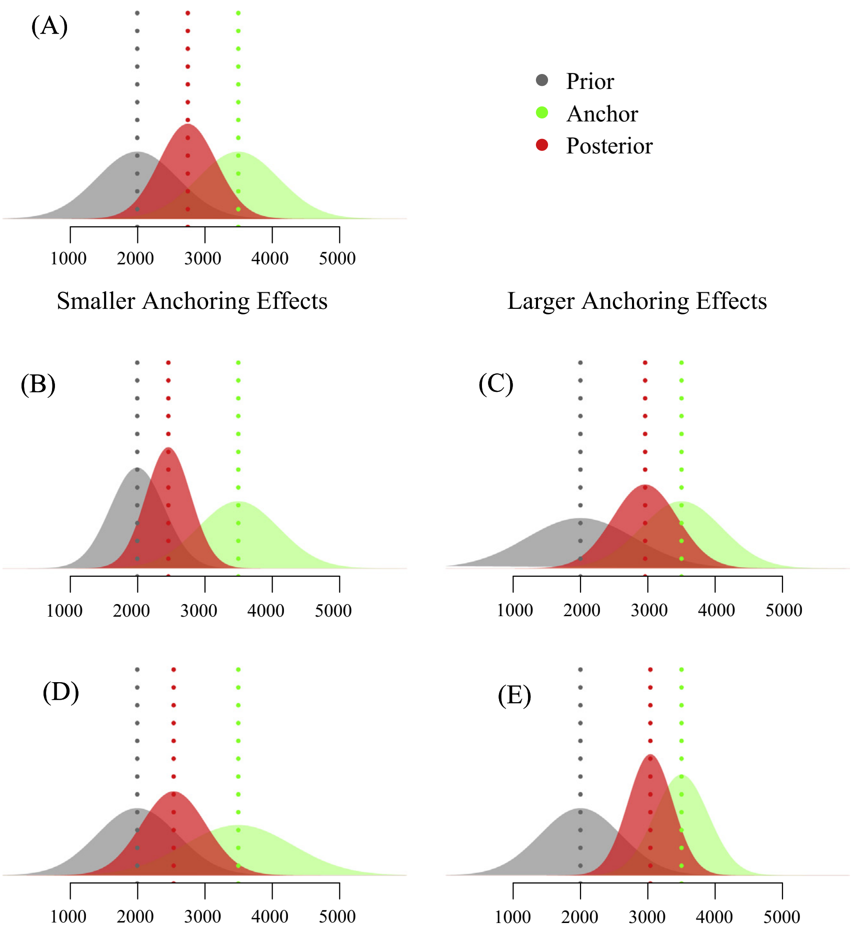

Fig. 1. Illustrative example of how anchor information and prior information are integrated when updating the posterior representation of a target item. Compared to Panel A, Panels B and D have a smaller proportion of anchor information to prior information resulting in a smaller influence of anchor information on the posterior representation (i.e., a smaller anchoring effect). Compared to Panel A, Panels C and E have a larger proportion of anchor information to prior information resulting in a larger influence of anchor information on the posterior representation (i.e., a larger anchoring effect).

vide more or less information, which is captured in the variance of the green information density functions. If an anchor provides relatively less information, then it will have a more diffuse representation (e.g., Fig. 1D). If an anchor provides relatively more information, then it will have a less diffuse representation (e.g., Fig. 1E).

AIM models the integration of anchor information with prior information to form a posterior representation of the target item. In the model derivation sections below we formalize this integration. In the meanwhile, to offer some insight into the information integration consider again Fig. 1. The posterior representation is an amalgamation of the prior and anchor information. We define anchoring effects as the shift from the prior to the posterior representation.[3] Within the AIM framework, anchor-

[3] Note that this is a slightly different conceptualization compared to the past literature. Traditionally, anchoring has been studied as the differences in judgments between low and high anchors (e.g., Tversky & Kahneman, 1974). In the current article, we characterize anchoring as the difference between the observed judgment in the presence of an anchor relative to what we would have expected their judgment to have been without the presentation of an anchor (i.e., the difference between the prior and posterior information density functions in Fig. 1).

ing effects are determined by the relative amount of prior information to anchor information. This intuition is partly informed by the observation that less informative anchors produce smaller anchoring effects (e.g., Wilson et al., 1996) and also by the observation that greater prior knowledge attenuates anchoring effects (e.g., Smith et al., 2013; Smith & Windschit, 2015). Thus, compared to Fig. 1A, smaller anchoring effects occur when individuals have more prior information (Fig. 1B) or less informative anchors (Fig. 1D). Similarly, larger anchoring effects occur when individuals have less prior information (Fig. 1C) or more informative anchors (Fig. 1E).

The above example is intended as a useful illustration for how AIM conceptualizes the prior and anchor information and their integration in the posterior representation of the target judgment. In the following sections we formalize this intuition and derive, first a version of AIM intended for between-participant data, then a version of AIM intended for within-participant data.

3. A between-participants version of AIM

We now present the explicit details of AIM so that it can be fit to empirical data. There are three major components to the model. First, we assume that for every possible stimulus, individuals have some representation of the target of interest. For the purposes of this article, the stimuli consist of general knowledge questions such as "What is the length of the Mississippi River?" The true solution to these stimuli is called the target, and we denote it θ_0. AIM assumes that each individual has a particular representation of what they believe to be the true state of θ_0, prior to the presentation of the anchor. We refer to this representation of θ_0 as the prior representation, and its *support* is denoted Θ (i.e., the set of θ values that have some probability of being chosen). The second component of AIM quantifies the amount of *perceived* information the anchor value α provides about θ_0, and we refer to this component as the information function. When presented an anchor value α, AIM assumes that individuals combine their prior representation with the information provided by the anchor to form a new representation about the target θ_0. This new representation is the third component of AIM, and we refer to it as the posterior representation. We will now describe each of these components in turn.

3.1. The prior representation

The first component of our model is the prior representation. The prior representation serves as the observer's internal representation of θ_0, which we will temporarily denote $\text{Prior}(\Theta)$. The prior representation is a function over Θ that provides the probability that each value of Θ is equal to the true state θ_0 (i.e., the density of information at any given value). This representation could take any form. It could be a single point on the continuum of possible values (e.g., $\text{Prior}(\Theta) = I(\theta = 2320)$, where $I(x)$ denotes an indicator function returning a 1 if x is satisfied and a 0 if it is not), or it could be a distribution over an entire continuum (e.g., $\text{Prior}(\Theta) = \text{Uniform}[0, 10,000]$). The information an observer has about θ_0 is twofold. First, the prior representation provides information about the most plausible value for θ_0, which can be determined by finding the point along the continuum Θ that corresponds to the highest density in the prior representation. Second, the prior representation quantifies the amount of uncertainty in the individual's information about θ_0. Prior representations with larger variances will reflect larger degrees of uncertainty whereas narrower prior representations will reflect a strong tendency toward a particular range along the continuum Θ (i.e., indicates greater prior information about the target item).

For this initial instantiation of AIM, we assume a convenient form for the prior representation. Specifically, we assume that the prior representation can be well approximated by a normal distribution with mean parameter θ and standard deviation parameter δ. To justify this assumption, it will sometimes be necessary to transform the space of Θ so that a normal distribution is appropriate. Because the prior representation will depend on both θ and δ, we will henceforth make this explicit in the notation by denoting the prior representation as $\text{Prior}(\Theta|\theta, \delta)$.

Although an assumption of normality may be a naïve one, we make it here for reasons of simplicity, interpretability, convenience, and psychological plausibility. First, the normal distribution easily quantifies a "best estimate" in the form of the mean parameter θ. Second, the normal distribution also provides a parameter for the dispersion of the representation δ, which reflects the uncertainty about θ_0. Larger values of δ will spread the prior over a much larger region of Θ, reflecting larger uncertainty about the target item. In addition, the choice of a normal distribution offers psychological plausibility by applying a symmetric, exponentially decreasing weight to values as a function of their distance from the mean θ – a common assumption in cognitive modeling with respect to both time and similarity (e.g., Kruschke, 1992; Lee, 2004; Lee & Wagenmakers, 2010; Liu & Aitkin, 2008; Nosofsky, 1986; Rubin & Wenzel, 1996; Shepard, 1987; Treisman & Williams, 1984; Turner et al., 2011; Wixted, 1990). Whereas the assumption of a convenient distribution works well for the current data, we advise that other more complicated prior representations may be necessary for some data (e.g., Lewandowsky, Griffiths, & Kalish, 2009).[4]

3.2. The influence of the anchor

We now turn to the quantification of the influence of the anchor. Although there are a myriad of supporting materials to justify the use of a normal prior representation, there are far fewer to aid us in selecting how the anchor might influence an individual's estimate. We draw from assumptions commonly made in categorization and similarity models. These models assume that as stimuli are presented, they activate internal representations of categories as a function of their distance from the perceived stimuli (e.g., Kruschke, 1992; Nosofsky, 1986). The further these perceptions are from the category representations, the less active the internal category becomes, which results in a decrease in the likelihood of that particular classification. Unlike categorization models, AIM assumes a continuous internal representation for the "classification" of the true state of θ_0. AIM assumes that the anchor possesses information about θ_0 that influences the internal representation. Clearly, some anchors will have a stronger influence on an individual's representation than others. To accommodate this feature, the information function can take many forms depending on the value of an "anchor influence" parameter λ. Because of the explicit dependence of the information function on the anchor α and the anchor influence parameter λ, we denote the information function $\text{Info}(\Theta|\alpha,\lambda)$.

We will again assume that the influence of the anchor has a convenient functional form. The form of the information function will take one of two distributions, depending on whether or not the anchor provides directional information. For anchors that do not provide directional information, we assume that the information function is normally distributed with mean equal to the anchor α and standard deviation equal to λ, such that

$$\text{Info}(\Theta|\alpha,\lambda) = \frac{1}{\sqrt{2\pi}\lambda}\exp\left[-\frac{(\Theta-\alpha)^2}{2\lambda^2}\right].$$

It is important to note that some anchors can contain directional information (Simmons, LeBoeuf, & Nelson, 2010). For example, whereas an anchor about the Mississippi river being 3500 miles long does not necessarily contain directional information, an anchor like "the length of the Mississippi river is less than 3500 miles long" provides specific information about the range of values of θ_0. We argue that directional and regular anchors have similar psychological properties but directional anchors will carry additional information provided by the item. If the anchor α carries with it a negative directional influence (e.g., less than 3500 miles long), then we assume that the information function will be normal in form but will have zero density in the region to the right of the anchor, such that

$$\text{Info}(\Theta|\alpha,\lambda) = \begin{cases} \frac{1}{\sqrt{2\pi}\lambda}\exp\left[-\frac{(\Theta-\alpha)^2}{2\lambda^2}\right] & \Theta \leqslant \alpha \\ 0 & \Theta > \alpha. \end{cases}$$

Similar functions can be formed to accommodate positive directional influences by substituting Θ with $-\Theta$, and anchors that are negative in value by replacing α with $-\alpha$.

[4] We are currently investigating more complicated, infinite dimensional prior representations using nonparametric Bayesian alternatives.

AIM's assumptions about the influence of the anchor are made mostly out of convenience. Specifically, by using a normal information function concurrently with a normal prior representation, we can provide closed form expressions for the posterior representation, which is convenient for estimation purposes. Notice that for directional anchors, the information function is not a probability distribution (i.e., it does not integrate to one). However, as we show below, we do not require the information function to be a probability distribution due to the integration step used to derive the posterior representation. Specifying the information function in this way allows us to compare the parameter λ across a variety of anchor types (e.g., regular and directional).

3.3. The posterior representation

The final component of the model is the posterior representation. Similar to the prior representation, the posterior representation is a distribution over Θ, and reflects the underlying beliefs about θ_0 following the presentation of the anchor value α. The posterior representation serves as the final result of the integration of a priori information (i.e., the prior representation) and information provided by the anchor. There are many ways in which the prior representation could be combined with the information function to form this new posterior representation.[5] One convenient and normative way to combine two distributions is through Bayes' rule.

We denote the posterior representation as $\text{Post}(\Theta|\theta, \delta, \alpha, \lambda)$, and we obtain it by multiplying the prior representation $\text{Prior}(\Theta|\theta, \delta)$ by the information function $\text{Info}(\Theta|\alpha, \lambda)$ and marginalizing, such that

$$\text{Post}(\Theta|\theta, \delta, \alpha, \lambda) = \frac{\text{Prior}(\Theta|\theta, \delta)\text{Info}(\Theta|\alpha, \lambda)}{\int \text{Prior}(\Theta|\theta, \delta)\text{Info}(\Theta|\alpha, \lambda)d\Theta}. \quad (1)$$

We again note that this posterior representation is a probability density function that provides the probability that each Θ equals the individual's estimate of the true state θ_0.

The marginalization constant in the denominator of Eq. (1) can be difficult to evaluate. However, our assumptions about how the prior representation is transformed into the posterior representation allow us to derive analytic expressions for both the prior and posterior representations. We use the terms prior and posterior to refer to the representations available before and after the presentation of an anchor. That is, a prior representation is the information an individual would have available to make the target judgment when no anchor is present. A posterior representation is the information an individual has which includes the prior and anchor information. Within the standard anchoring paradigm, only the posterior representation is actually constructed to form a response (i.e., a judgment is not elicited before the presentation of the anchor). Therefore, the prior representation in our model represents the hypothetical prior representation that would have been constructed had a judgment been elicited without an anchor. Hence, for most applications, the prior representation is a latent construct to be inferred from the data.

3.3.1. Eliciting a response

Although AIM generally assumes that individuals have access to an entire (posterior) representation (i.e., not a single point), participants in an experiment are typically asked to form a response of a single value. As a consequence, to form a testable model we must specify an elicitation mechanism that converts the posterior representation $\text{Post}(\Theta|\theta, \delta, \alpha, \lambda)$ to the participant's estimate $\widehat{\Theta}$. When the estimate $\widehat{\Theta}$ is a single value, we could simply adopt the rule $\max_\Theta\{\text{Post}(\Theta|\theta, \delta, \alpha, \lambda)\} = \widehat{\Theta}$. However, this particular rule may not be what the individual had intended. For example, this rule may not be the best choice if $\text{Post}(\Theta|\theta, \delta, \alpha, \lambda)$ is a bi-modal or uniform distribution.

Rather than assuming that an individual always produces an estimate at the mode of their representation, AIM assumes that individuals sample evidence probabilistically from their posterior representation. For example, when an individual provides an estimate, they randomly sample a single value from their posterior representation, and as a consequence, they are more likely to sample from higher-

[5] Indeed, the combination of these two sources of information may vary across tasks, context, anchor types, and individuals. However, for our purposes, we will not allow the combination process to vary.

density regions. We assume random sampling for convenience, but acknowledge that individuals more frequently respond with rounded numbers. Although the actual process may be a pseudo-random sampling of the posterior representation, we assume random sampling within the model for simplicity.

The number of samples n may be treated as another parameter in the model, and it performs similar to other "threshold" parameters common in models of evidence accumulation, such as models of response time (e.g., Brown & Heathcote, 2005, 2008; Ratcliff, 1978; Usher & McClelland, 2001). In this article, we will not explore the role n plays and will instead assume that $n = 1$ throughout. This assumption is justified by other recent results suggesting that individual form judgments on the basis of only a few noisy samples from their (posterior) representation (e.g., Griffiths & Tenenbaum, 2006; Vul, Goodman, Griffiths, & Tenenbaum, 2014). Hence, we assume a single sample of evidence is drawn from the posterior representation, and this sample serves as the participant's best estimate for θ_0.

3.4. Summary

AIM has three parameters that control how each source of information is used in generating the observed data. Fig. 1 shows some hypothetical model predictions under a few settings of the model parameters. In all panels, the mean of the prior representation θ is held constant to exhibit the influence of the other parameters δ and λ. The prior representation, anchor information function, and posterior representation are illustrated by the gray, green, and red density functions, respectively. For example, Fig. 1B and C illustrate the effects of the parameter δ on the posterior representation. Under strong prior information (i.e., a small value of δ), the prior representation is less influenced by the anchor information, and as a result, the posterior representation is nearer to the prior. However, under weak prior information (i.e., a large value of δ), the prior representation is susceptible to the influence of the anchor, and so the posterior representation is nearer to the anchor value.

Fig. 1D and E show the influence of the anchor influence parameter λ under weak and strong influence settings, respectively. Under weak anchor information (i.e., large values of λ; Fig. 1D), the prior representation is less influenced, resulting in only a small shift from the prior representation to the posterior representation. Under strong anchor information (i.e., small values of λ; Fig. 1E), the prior representation is highly influenced, resulting in a relatively large shift from the prior representation to the posterior representation.

Having discussed all the major components of AIM, we can now turn to evaluating the model on the basis of empirical data. Experiment 1 was designed to demonstrate AIM's ability to describe the effects of an assortment of anchor values. We first describe the details of the experiment, and all of the essential effects present in the data. We then discuss the implementation details of the between-participants version of AIM and fit it to data. We close the next section by discussing the modeling results.

4. Experiment 1

Experiment 1 was intended to test the model across a range of judgment targets and a wide variety of anchoring effects. The following sections will demonstrate the etiologies of regular anchoring effects, directional anchors, irrelevant anchors, and the influence of cognitive load on anchoring effects.

4.1. Methods

Participants were recruited from Amazon's Mechanical Turk website (see Buhrmester, Kwang, & Gosling, 2011; Paolacci, Chandler, & Ipeirotis, 2010, for a review of Mechanical Turk as a viable source of experimental data) and completed the task in exchange for $0.50. A total of 589 participants (sub sample information will be provided by condition) completed one of the five between-participant conditions: prior, regular anchors, directional anchors, irrelevant anchors, and cognitive load.

4.1.1. Materials

Participants in every condition were presented the same 6 anchoring problems taken from Simmons et al. (2010) and Jacowitz and Kahneman (1995). The six target questions were as follows: (1) the length of the Mississippi River, (2) the average annual rainfall in Philadelphia, (3) the maximum speed of a house cat, (4) the average annual temperature of Phoenix, (5) the population of Chicago, and (6) the height of the tallest redwood tree. Once the correct answer was determined, for some questions we set anchors that were either higher or lower than the true value. Questions 1, 3, and 5 used high anchors (3500 miles, 55 miles per hour, and 5 million people, respectively) whereas Questions 2, 4, and 6 used low anchors (25 in., 52 °F, and 65 feet, respectively).

4.1.2. Procedure

The five between-participant conditions are outlined in Table 1. In all conditions participants estimated each of the six target items. Thus, Experiment 1 employed a 5 between-participant (Anchoring Condition) × 6 within-participant (Target Questions) mixed design. In addition to the prior (i.e., unanchored control condition) and regular anchor conditions, we included three additional anchoring paradigms that have received notable attention within the anchoring literature. Because anchors often provide directional information we included a condition with directional anchors (Simmons et al., 2010). The literature on irrelevant anchors and numeric primes has been controversial to say the least, with evidence toward (Adaval & Monroe, 2002; Critcher & Gilovich, 2008; Englich, 2008; Oppenheimer et al., 2008; Mussweiler & Englich, 2005; Wilson et al., 1996; Wong & Kwong, 2000) and against (Brewer & Chapman, 2002; Chapman & Johnson, 1994; Frederick & Mochon, 2012; Mochon & Frederick, 2013; Mussweiler & Strack, 2001) such effects. As such, we included a condition where the anchors were in reference to items different than those of the target judgments. Lastly, given the research that used cognitive load as a means of studying anchoring effects (Blankenship, Wegener, Petty, Detweiler-Bedell, & Macy, 2008; Epley & Gilovich, 2001, 2006; van Rooijen & Daamen, 2006), we included a condition where participants completed the regular anchoring task under cognitive load. We included this range of conditions to test AIM's ability to model a range of anchoring paradigms.

Borrowing terminology from above, we refer to estimates made by participants in the prior condition as prior estimates and estimates made by participants in the anchoring conditions as posterior estimates. Roughly 3% of the data were deleted due to either extremely implausible answers or because they attempted to look up the answers on-line. We were able to identify the most accessible answers through Wikipedia, Google, and Yahoo searches. If participants had longer response times (e.g., greater than 10 s) and answers identical to the ones we identified through these websites, those observations were deleted. Also as part of the 3% of deletions, estimates in the directional anchoring condition that were on the incorrect side of the anchor (e.g., an estimate of 4500 miles to an anchor that was stated to be less than 3500 miles) were deleted. This method of deletions is consistent with Epley and Gilovich (2001, 2006) and Simmons et al. (2010).

4.2. The model

To model the data from Experiment 1, we used the between-participant version of AIM described above. We assumed separate prior representations for each of the six questions, and importantly, this prior representation was common to all participants. For each question, there were a total of four

Table 1
Outline of between-participant conditions for Experiment 1. The example question that each description below corresponds to is "What is the length of the Mississippi River?" Key: CL is Cognitive Load.

Condition	Description/example
Prior	Unanchored judgment (i.e., no additional information presented)
Regular	Is the Mississippi River greater than or less than 3500 miles?
Directional	The Mississippi River is less than 3500 miles
Irrelevant	Is the distance between Cairo and Istanbul greater than or less than 3500 miles?
CL	Memorized 9-letter string before anchor then provided string after the judgment

anchoring conditions. To model the effects of these four types of anchors, we assumed separate λ parameters for each question by anchor type, and these λ parameters were assumed to be common to all participants. Although it would be convenient to have a single estimate for λ corresponding to each anchoring condition across questions, the questions are on different scales making it unclear how to constrain the information function to incorporate this idea. Thus, when fitting the model to the data, we partitioned the data by question, so that the analysis was independent across questions. In the analysis of each question, we avoided building a hierarchical structure on the λ_ks because it was not clear what distribution the λ_ks might take, and we wished to avoid any unjustified (and unnecessary) restrictive assumptions. Hence, a total of six θs, six δs, and 24 λs were used when fitting the model to data.

Fig. 2 displays a graphical diagram of the model used in each question. These diagrams are often used in cognitive modeling because they convey the specification of a mathematical model in a convenient path diagram (e.g., Lee, 2008; Lee & Wagenmakers, 2010). Each "node" in Fig. 2 corresponds to a separate variable in the model, and shaded nodes represent variables that are known, such as the data x or the value of the anchor α_k. Nodes that are connected to one another indicate that the variables in the nodes are dependent. This dependence travels one way, such that the "pointing" node controls the variable it points to. For example, Fig. 2 shows the dependence of the data y_k on all three parameters of the model, θ, δ, and λ_k, which is expressed mathematically in Eqs. (A.6) and (A.7). Finally, graphical diagrams represent vector-valued dependencies among variables by surrounding these variables by a rectangle, which is commonly referred to as a "plate." In Fig. 2, the variables λ_k, α_k, and y_k are all on the plate because they each contain four elements – one element for each (posterior) Condition k.

We fit the model to the data in a Bayesian fashion, which required the specification of prior distributions for each parameter. For ease of presentation, we have relegated the technical details of the model to Appendix A, and we have also provided JAGS code for fitting the between-participant version of AIM to data in Appendix C.

4.3. Results

To fit the model to the data in a Bayesian framework, we implemented a Differential Evolution Markov chain Monte Carlo (DE-MCMC) sampler with 24 chains (see ter Braak, 2006; Turner, Sederberg, Brown, & Steyvers, 2013, for details). We have also provided JAGS (Plummer, 2003) code for performing the estimation in Appendix C. We ran the chains in parallel for 10,000 iterations and discarded the first 2000 iterations as a burnin, resulting in a total of 192,000 samples from the posterior distribution. Standard methods were used to ensure convergence.

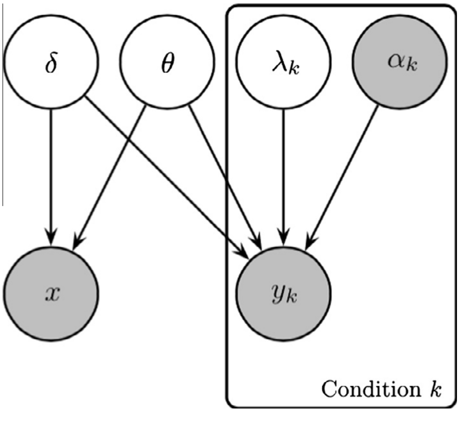

Fig. 2. A graphical diagram of the between-participant version of AIM. Nodes with explicit dependencies are connected with arrows, and plates are used to indicate a loop over an index. Nodes corresponding to fixed quantities, such as data or the anchor value, are shaded light gray, whereas unknown parameters are not shaded.

Fig. 3 shows a violin plot of the posterior distributions for each of the λ_k parameters on the log scale. Each panel in Fig. 3 shows the estimates for λ for each anchor type (*x*-axis) for a particular question. Because λ is the standard deviation of a normal distribution, as λ increases, the influence of the anchor diminishes and the prior representation is more reflective of the posterior representation. Across questions, the estimates for λ have little meaning, but within a question, the relative differences in the estimates convey the perceived importance of the anchor. For example, across questions we see that the estimates for λ in the irrelevant anchor condition are larger than the estimates in other conditions. This result demonstrates that AIM is capturing patterns in the data that conform with the a priori hypothesis that irrelevant anchors should produce smaller anchoring effects. The pattern of results for λ estimates in other conditions is less clear. For example, λ in the directional condition is sometimes lower (e.g., Questions 1, 2, and 3), and sometimes larger (e.g., Questions 4, 5, and 6) than λ in the regular condition.

We can also examine the posterior distributions for the parameters θ and δ corresponding to the prior representation. Table 2 shows the median and 95% Bayesian credible set of the estimated posterior distributions for θ and δ, separated by question. For example, on Question 1 (The length of the Mississippi River), the mean estimate of θ was $6.916 = \log(1008.279)$ and the estimated uncertainty parameter δ was 0.858 on the log scale. Converting these parameters to a representation on the normal scale produces a 95% credible set of (185.92, 5363.22). As with the parameter λ, it is difficult to draw comparisons across items for these parameters because the range of their support is different.

The final analysis we preformed was to generate model predictions for both the prior and posterior representations used by the subjects in this experiment. To do this, we randomly sampled 10,000 observations from the joint posterior distribution of all model parameters, and used these values to construct both the prior and posterior representations from the model. The resulting distributions are known as posterior predictive distributions, and are plotted in Figs. 4–7 as gray lines against the density of the data (black lines). Each panel in these figures corresponds to a different question used in the experiment. The black dashed lines represent the observed data from the prior condition, and is reproduced in each of the four figures for convenience (e.g., dashed black line in the top left

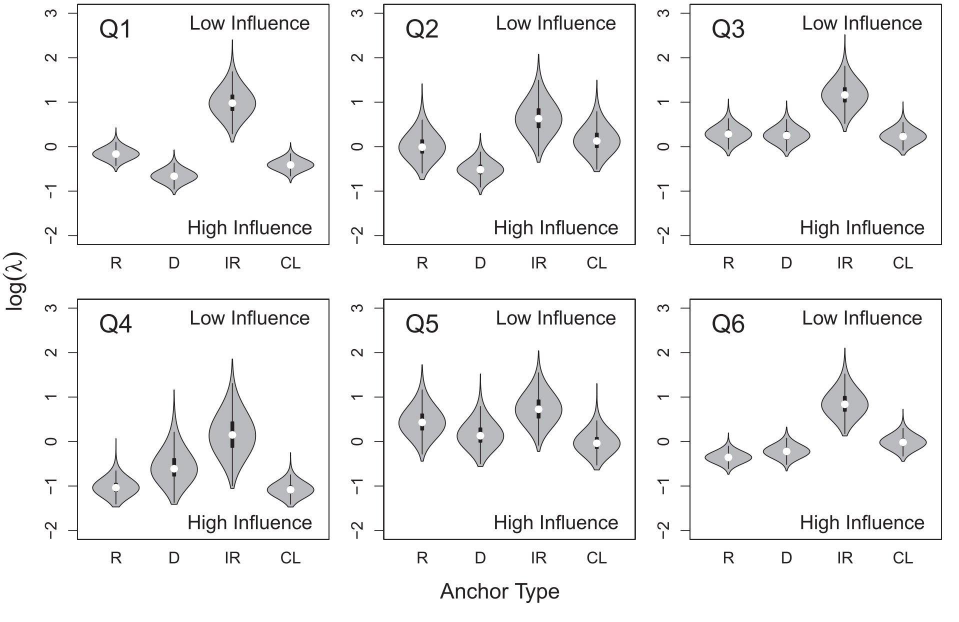

Fig. 3. The estimated posterior distribution of λ_k for each of the six anchoring conditions for each question. From left to right, the anchor types are: regular (R), directional (D), irrelevant (IR), and cognitive load (CL). The parameter λ_k denotes the degree of anchor influence, where areas of high anchor influence are small values and areas of low anchor influence are larger values.

Table 2
Summaries of the estimated posterior distributions of θ and δ obtained for each item in Experiment 1.

Question	$\widehat{\theta}$		$\widehat{\delta}$	
	Median	Credible set	Median	Credible set
1	6.916	(6.834, 6.999)	0.858	(0.811, 0.910)
2	3.287	(3.228, 3.345)	0.738	(0.690, 0.788)
3	2.526	(2.454, 2.599)	0.791	(0.750, 0.833)
4	4.427	(4.413, 4.440)	0.147	(0.140, 0.154)
5	1.316	(1.256, 1.377)	0.767	(0.720, 0.817)
6	5.244	(5.166, 5.320)	0.807	(0.764, 0.854)

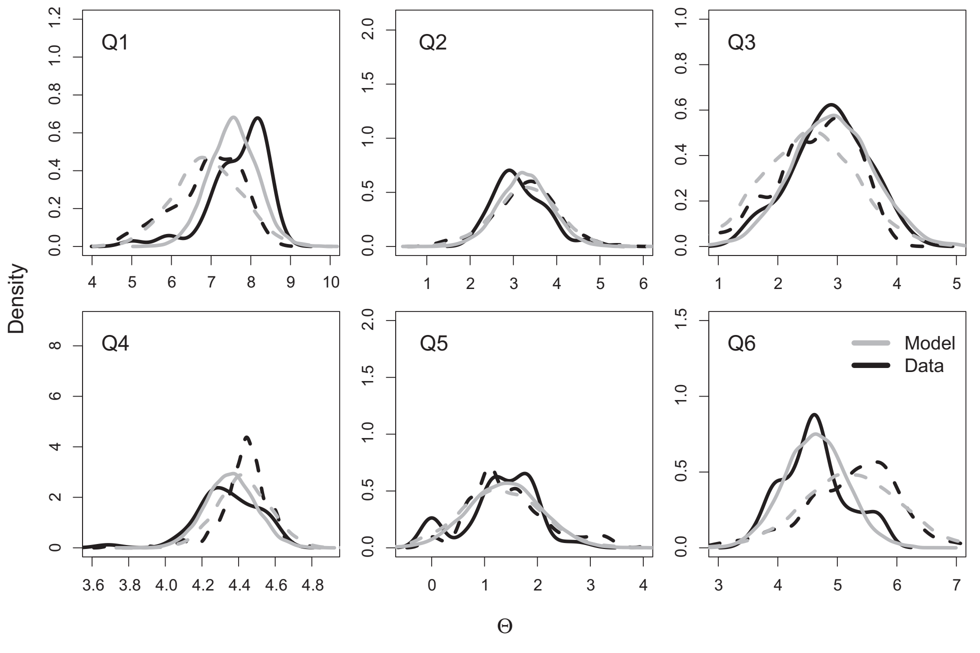

Fig. 4. The distribution of the log-estimates of Θ for each question in the regular anchors condition (solid black lines) along with the distribution of the log-estimates in the prior condition (dashed black lines). The corresponding posterior predictive densities obtained after fitting the model are shown by the gray lines (prior representations are dashed whereas the posterior representations are solid).

panel of Fig. 4 is the same as the dashed black line in the top left panel of Fig. 5). In general, AIM closely matches the observed data across all conditions and questions. In conditions where the posterior did not differ appreciably from the prior (e.g., the irrelevant anchoring condition) AIM predicts a similar pattern and is less influenced by the anchor value. By contrast, in conditions where the anchor had a remarkable influence on the prior, such as in the directional condition, our model peaks in the same regions as the data indicating a similar interaction with the provided anchor. In conditions where the model tends to provide a relatively poorer fit to the data, the normality assumption about the prior distribution tends to be invalid. In these conditions, if one assumed a more complicated function for the prior representation or the information function, one could obtain better fits to the data.[6] We will now discuss the results in more depth for each type of anchor in turn.

[6] For example, one could assume a Dirichlet process or finite mixture distribution for the prior representation (or information function or both), and obtain much better fits.

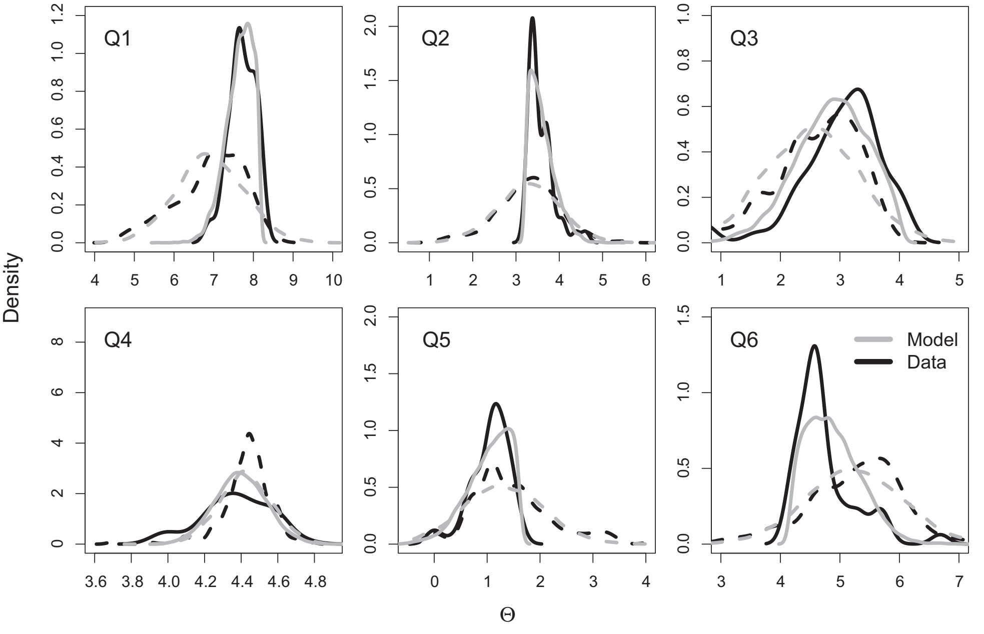

Fig. 5. The distribution of the log-estimates for each question in the directional anchors condition (solid black lines) along with the distribution of the log-estimates in the prior condition (dashed black lines). The corresponding posterior predictive densities obtained after fitting the model are shown by the gray lines (prior representations are dashed whereas the posterior representations are solid).

4.3.1. The prior condition

To get an accurate approximation of the aggregate prior representation of individuals in the anchoring conditions, 323 participants (61% female with a mean age of 35) completed the task without the presentation of any anchors. Experiment 1 hinges on the assumption that the distribution of estimates in the control (i.e., prior) condition were representative of the prior representations of participants in the anchoring conditions.

The prior distributions for each condition appear in Figs. 4–7 as the dashed black lines. Because these priors are from a separate condition of the experiment and do not depend on the type of anchor presented, the priors shown in each of these figures are equivalent, and are only shown for interpretation purposes for the other experimental conditions. That is, because the magnitude of the shift from prior to posterior representation depends on the anchor type, we can interpret the effects of the anchor in each condition by examining the discrepancy between the solid and dashed curves in each of these four figures.

The model fits to the prior condition are shown as the dashed gray densities in Figs. 4–7. Generally speaking, the model captures the shape of the prior representations well, despite the noise present in the data. To better account for this noise, one could increase the flexibility of the prior representation. However, our Gaussian prior representation assumption provided an adequate fit for our data, and so we save a thorough investigation of these more complex functional forms for future research.

4.3.2. Regular anchors

Sixty participants (56% female with a mean age of 33) answered a comparative question (e.g., is the Mississippi River greater than or less than 3500 miles long?) about whether the target was greater than or less than the anchor value we provided, before making their estimate, for each of the six questions. The distribution of participants' responses for each of the six questions are presented in Fig. 4. As mentioned above, the black dashed lines represent the distribution of estimates provided by partici-

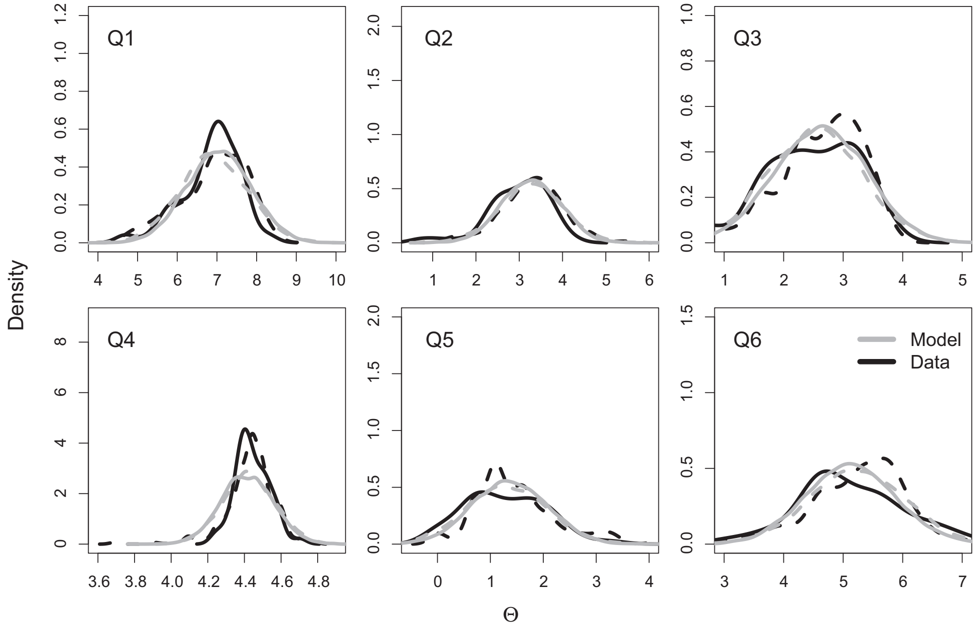

Fig. 6. The distribution of the log-estimates for each question in the irrelevant anchors condition (solid black lines) along with the distribution of the estimates in the prior condition (dashed black lines). The corresponding posterior predictive densities obtained after fitting the model are shown by the gray lines (prior representations are dashed whereas the posterior representations are solid).

pants in the prior condition, whereas the black solid lines are the distributions of the estimates provided by the participants in the regular anchoring condition. The solid gray lines represent the model predictions for the regular anchoring condition. In general, the regular anchor does provide some noticeable shift in the posterior representation relative to the prior representation. This shift is most pronounced for Question 6, and is also noticeable for Question 1. The model captures these shifts well: for Questions 1 and 6, the model predicts a shift of a similar magnitude and in the same direction as the empirical data, whereas for the remaining questions, the shift is negligible in a way that matches the observed data.

4.3.3. Directional anchors

Seventy-two participants (65% female with a mean age of 34) were presented an anchor statement (e.g., the Mississippi River is less than 3500 miles long) before producing an absolute judgment. Fig. 5 shows the density of the posterior distributions for each question in the directional anchoring condition as solid black lines. The model predictions are again shown as the solid gray lines. Fig. 5 shows that the effects of the directional anchors are very strong, producing large differences between the prior and posterior representations. In particular, for all of the questions except Question 4, the posterior representation is more peaked, having less variance relative to the prior representation. This pattern reflects an increase in the amount of available information in forming the judgment following the presentation of the anchor, and suggests that participants do use the information provided by the anchor when eliciting their response.

Generally speaking, Fig. 5 shows the model captures the anchor effects well, producing marked changes in the posterior representation that closely match the changes observed in the data. The fits are particularly good for Questions 1, 2 and 5, whereas some misfit occurs for Questions 3, 4, and 6. The reason for these misfits is mixed. For example, in Question 3, the misfit of the posterior representation seems to be due to a misestimation of the prior representation, whereas in Question 6, the mis-

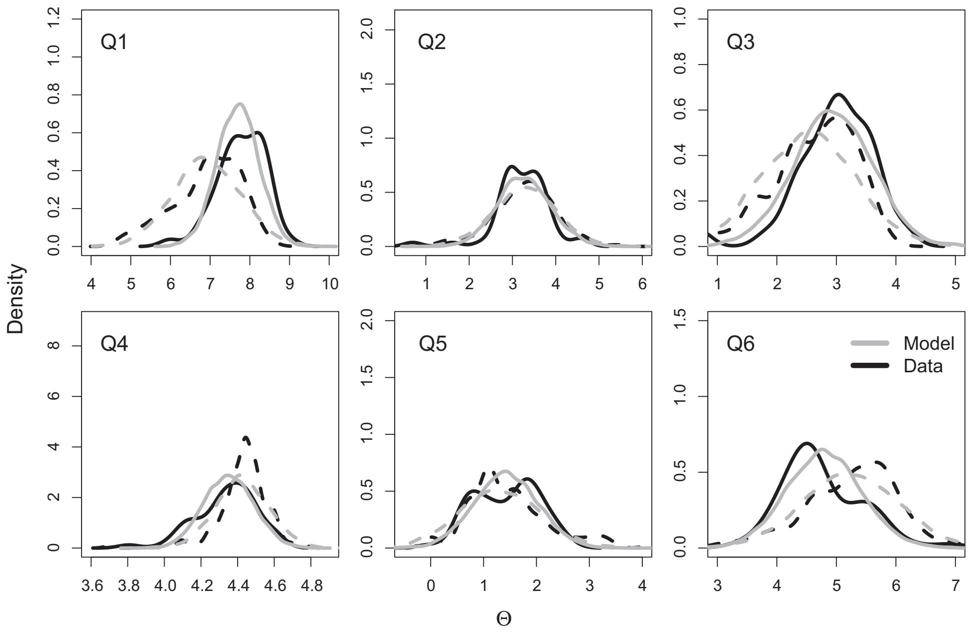

Fig. 7. The distribution of the log-estimates for each question in the cognitive-load condition (solid black lines) along with the distribution of the log-estimates in the prior condition (dashed black lines). The corresponding posterior predictive densities obtained after fitting the model are shown by the gray lines (prior representations are dashed whereas the posterior representations are solid).

fit seems to be due to a misestimation of the anchor influence. We suspect that much of this misestimation is due to irregularities in the representations themselves, namely that these representations are not Gaussian.

4.3.4. Irrelevant anchors

Seventy participants (67% female with a mean age of 37) were presented the six anchoring questions with irrelevant comparison questions (e.g., is the distance between Cairo and Istanbul greater than or less than 3500 miles?). Irrelevant anchors should not provide much information when forming a posterior representation, and so we would expect the estimates for λ in this condition to be large. However, the proportion of information provided by the anchor will be larger when an individual has little prior information (i.e., when δ is large). Hence, we might generally expect λ and δ to be negatively correlated.

Fig. 6 shows the participants' estimates from the prior condition (dashed) against the posterior estimates of participants in the irrelevant anchor condition (solid) for each of the six questions. Fig. 6 shows that the anchoring effects in this condition are relatively weak because the posterior representations are highly similar to the prior representations. The model predictions are again shown as the gray lines. Across the panels, Fig. 6 shows that the model predictions closely match the patterns in the observed data, but this is unsurprising given that the effects are not particularly strong.

4.3.5. Cognitive load

To develop a better understanding of the influence of cognitive load on anchoring effects, 64 participants (56% female with a mean age of 33) completed an identical anchoring task to that of the regular anchor condition. Just before the comparative question was presented with the anchor, participants were shown a nine letter string and were told to memorize it for later recall. For each question, after producing an absolute estimate of the target item, participants were asked to recall the series of letters. Fig. 7 plots participants' estimates from the prior condition against posteriors in the cognitive-load condition for each of the six questions. The effects of the cognitive-load condition were relatively strong in Questions 1 and 6, and to a lesser extent in Question 3. The model predictions are again shown as the gray lines, where solid lines represent posterior conditions and dashed lines represent the prior condition. Fig. 7 shows that the model again matches the basic patterns in the data well. The match to the data is particularly close in Questions 1 and 6, where the effects of the anchor are strong.

4.3.6. Summary of experiment 1

Experiment 1 demonstrated that the between-participant version of AIM can account for a wide variety of anchors and moderators. AIM generally produced predictions that closely matched the observed data across six questions and four types of anchoring effects. There were some mismatches in the predictions, but we suspect that the majority of these mismatches are due to the lack of complexity in the prior representation. We speculated that increasing the flexibility of this component of the model would improve the fitting results.

In the sections above, we have outlined a between-participants version of AIM that allows us to investigate the effects of different anchor types across groups of individuals. Overall, we would consider the fits of our model to be adequate, considering the extreme variability in participants' estimates and the lack of multiple observations per participant. Although this version of the model is very useful for understanding how the location and context of an anchor influences judgments, we have not provided a detailed modeling account of how anchors might influence individual judgments. In the next section, we extend AIM to account for individual-specific anchoring effects.

5. A within-participant version of AIM

The version of AIM outlined and tested above provides a cogent descriptive explanation of the influences of different types of anchors on participants' judgments. One shortcoming of the model is the assumption of the homogeneity of the anchor influence parameter across individuals (i.e., predicting anchoring effects at the group level rather than the individual level). In the next section we relax this assumption and extend AIM to account for anchoring effects at the individual level.

5.1. Prior uncertainty and the role of confidence

The data that we typically acquire from an anchoring experiment has only one judgment per question per individual. Within AIM it becomes difficult to estimate both the influence of an anchor and the uncertainty in the prior representation *simultaneously*. Thus, to investigate individual differences with such little information, it is essential that we either apply reasonable constraints to the model or obtain more information about the judgment. For the prior uncertainty parameter, we will build both of these methods into AIM.

In our Experiment 2 (discussed below), we will obtain confidence judgments in addition to the estimates of θ_0. We disentangle the role of confidence by eliciting both prior and posterior judgments within a participant, where their confidence response to the prior judgment represents the prior uncertainty (i.e., how much prior information they have about the target item). We suspect that as the confidence in the judgment increases, the prior uncertainty parameter should decrease. However, we make no restrictive assumptions about the functional form of this relationship. Instead, we will use a separate prior uncertainty parameter for each confidence response option (e.g., six options if the scale is 1–6). Thus, we can estimate the prior uncertainty by assuming that this parameter is fixed either across different participants or across repeated anchors on the same continuum (i.e., multiple judgments about the same target item).

We agree that in some cases, it is unwise to assume that confidence judgments have the same meaning, even when they are all about the same target item. For example, it seems unlikely that when Participant A elicits a confidence judgment of "1" or "Very Unconfident" that this degree of confidence will have the same meaning for Participant B, who has also elicited a confidence judgment of "1". Even though these two participants have expressed the same degree of confidence (but may have different prior judgments about a target), their level of *uncertainty* may not be equivalent.

While the assumption above may be an overly restrictive one, we emphasize that this assumption is not necessary for our model. For example, with multiple confidence judgments per individual, one can assign a separate prior uncertainty parameter for each category, even when the target dimension Θ_k is not the same. To do so, we only require that the confidence judgments be related to one another in some meaningful way. For example, one could assume that the different confidence judgments come from the same overarching process, so that when a confidence judgment of "1" is provided for the estimate in Question k, it will be related to a confidence judgment of "1" in Question l. As another example, one could assume a functional form (e.g., a linear relationship) for how the confidence judgments are related to the prior uncertainty. As a result, having information about other confidence judgments such as "1" and "4" will inform our estimates of confidence judgments such as "3", even if this judgment was never actually elicited in the experiment.

In an attempt to create a more parsimonious model, we now discuss a way to relate the influence that a given anchor value may have to the anchor's consistency with an individual's prior representation.

5.2. Prior consistency

In order to make testable predictions about how an anchor might influence a judgment, we assume that participants use a measure of "prior consistency" to dictate how their prior representations should be updated. That is, consistency refers to the extent that the anchor information and the prior representation overlap. Given the prior representation $\text{Prior}(\Theta|\theta, \delta)$ and the presentation of an anchor value α, we assume that participants determine how likely the candidate solutions Θ are to be less than the presented anchor value α, under their prior beliefs. To do so, participants integrate their prior representation across the space $(-\infty, \alpha)$, or

$$\psi = P(\Theta \leqslant \alpha) = \int_{-\infty}^{\alpha} \text{Prior}(\Theta|\theta, \delta) d\Theta, \tag{2}$$

which is the cumulative distribution function under a participant's prior beliefs, and $\psi \in [0, 1]$ because it is a probability measure.[7]

The prior consistency variable ψ now provides a direct evaluation of how consistent a given anchor value is with a participant's prior beliefs. By inspecting various anchors under a fixed prior representation, we can see how the degree of consistency is non-monotonic. Specifically, as ψ approaches 0.5 (from either direction), the degree of consistency increases, but as ψ moves away from 0.5 (i.e., toward 0.0 or 1.0), the degree of consistency decreases. That is, anchor values that are much lower than the prior representation along the axis Θ will be very inconsistent, which will make ψ very small. By contrast, anchor values that are much larger than the prior representation along Θ will also be very inconsistent, but in this case, ψ will be large because the area under the prior representation will be larger. As a concrete example, consider a prior representation where $\theta = 0$ and $\delta = 1$ and the presentation of three different anchor values; $\alpha = \{-3, -0.5, 1\}$. The top left panel of Fig. 8 shows the model under these parameter settings where the solid black line (A1), dashed light gray line (A2) and dashed dark gray line (A2) correspond to the three anchor values respectively. The top right panel of Fig. 8 shows the prior consistency values $\psi = \{0.001, 0.309, 0.841\}$ for each anchor with the corresponding vertical lines. The most extreme anchor value $\alpha = -3$ has the smallest prior consistency ($\psi = 0.001$) and the largest anchor value $\alpha = 1$ has the largest ψ value ($\psi = 0.841$), but is still inconsistent with the prior representation.

Although the evaluation of prior consistency is essential for understanding how a given anchor might influence a participant's representation, we have not yet described how this relationship might unfold. We now discuss how the anchor influence parameter λ is determined by mapping ψ to λ.

[7] To be clear, we do not contend that participants actually compute an integral to evaluate the prior consistency. However, the calculation is a simple mathematical approximation to the actual process that might occur.

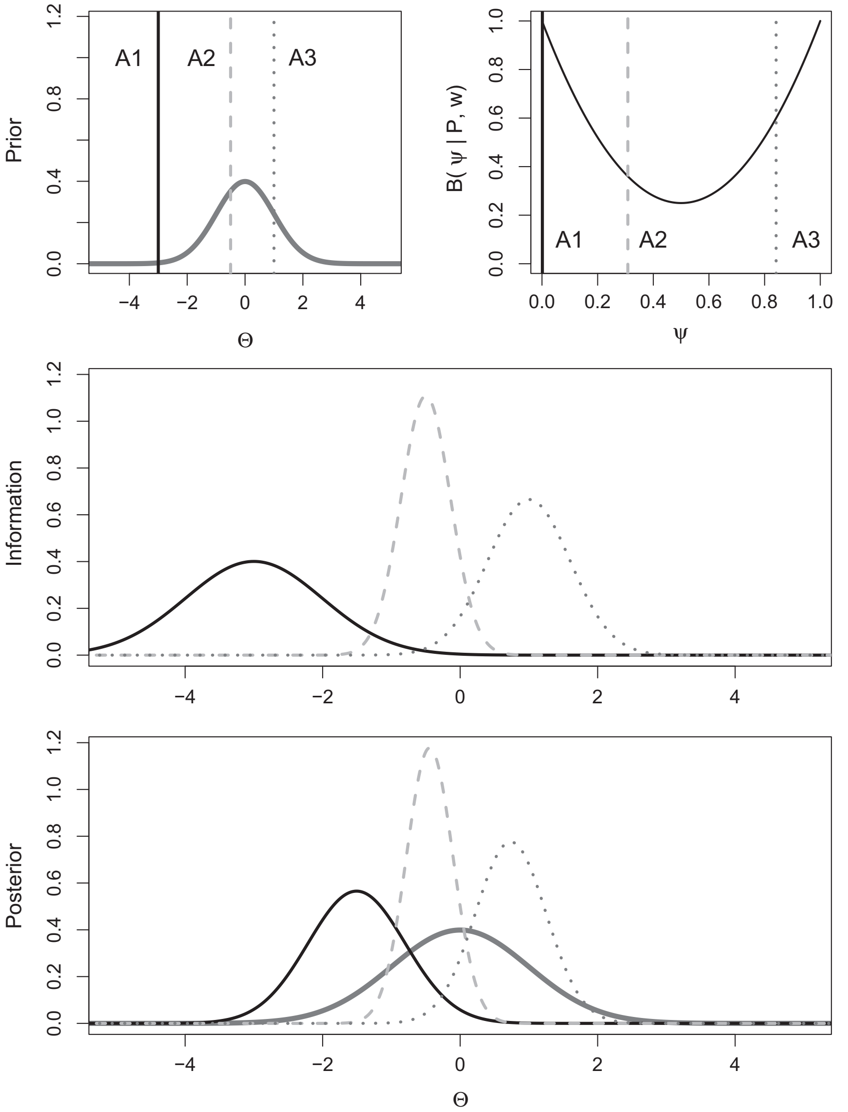

Fig. 8. The within-participant version of AIM. The top left panel shows a prior representation under the presentation of three different anchor values. The top right panel shows the mapping function relating the anchor influence parameter λ to the prior consistency variable ψ. The middle and bottom panels show the information functions and posterior representations corresponding to each anchor value, respectively.

Whereas λ represents the influence of the anchor for a group of participants in the previous version of the model, it now represents the influence of the anchor for a given individual.

5.3. Anchor influence

We assume that the influence of the anchor will be a direct function of the prior consistency variable ψ. However, the degree to which the prior consistency variable influences the prior representation is assumed to vary across participants.

5.3.1. Constraining the estimate of λ

To properly define a mapping function, we must first understand the behavior of the information function when λ becomes large. In particular, the information function has a very similar influence on the prior representation once λ increases beyond a certain value, which we will denote $\lambda_{\max}$. Once $\lambda > \lambda_{\max}$, the prior representation will be very similar to the posterior representation, and further increases in λ will not appreciably change the difference between the prior and posterior representations. This problem is not generally an issue for the previous version of our model because the effects of the anchor are pronounced for enough individuals that λ is much less than $\lambda_{\max}$. However, boundary problems do occur in the extension of our model because some individuals are completely insensitive to the presentation of the anchor, and as a result, we must build this behavior into the model.

Fortunately, a good selection of $\lambda_{\max}$ can be made prior to fitting the model. To do so, we must first examine the distribution of prior judgments across all individuals. Suppose after an experiment, we gather the set of prior judgments x about the target Θ prior to the presentation of an anchor located at α. To select an approximate value for $\lambda_{\max}$, we can use a nonparametric approach to first construct a density estimate $f(x)$ for x, and then assume that this density estimate is an adequate approximation, on average, of any given individual in the data set. We can then iteratively propose values $\lambda^*_{\max}$ for $\lambda_{\max}$ and examine the degree of difference between the prior representation $f(x)$ and the posterior representation, given by

$$\text{Post}(\Theta|\lambda^*_{\max}, \alpha) = \frac{f(x)\text{Info}(\Theta|\alpha, \lambda^*_{\max})}{\int f(x)\text{Info}(\Theta|\alpha, \lambda^*_{\max})d\Theta}.$$

To form an estimate for $\lambda_{\max}$, we slowly increase $\lambda^*_{\max}$ until the posterior representation does not appreciably differ from the prior representation. For example, the degree of difference between the posterior and prior representations could be calculated via the Kolmogorov-Smirnov or Kullback-Leibler test statistics. The smallest value of $\lambda^*_{\max}$ for which this condition holds will serve as the estimate for $\lambda_{\max}$.

It is worth mentioning that there are other, more principled ways of estimating $\lambda_{\max}$. For example, one could build the parameter directly into the model and estimate it on the group level in a Bayesian hierarchy. However, we have found that the quality of the model fit is fairly robust to the selection of $\lambda_{\max}$ as a result of the shrinkage that occurs across subjects due to the constraints imposed by the hierarchical model, and so more complicated estimation techniques may prove unnecessary.

5.3.2. The mapping function

We next define a mapping function to relate the prior consistency variable ψ to the influence of the anchor parameter λ. In theory, this function may be very complex, requiring many parameters to relate the two variables precisely for all participants. However, in practice, this function should posit psychologically meaningful parameters and will likely require some simplifying assumptions to facilitate the interpretability of these parameters.

For simplicity, AIM assumes that as the distance between the prior representation and the anchor increases, the anchor provides a marginally decreasing influence on the posterior representation, producing an inverted U-shaped relationship between anchor extremity and the magnitude of the shift from prior to posterior representation. This assumption is informed by the inverted U-shaped relationship observed by Wegener et al. (2001). That said, there are potential special cases where this assumption may not hold. First, if an individual uses an anchor as a hint (Schwarz, 1994; Wegener et al., 2010) it is feasible to assume that the hint is an exaggeration. If a speaker states that it was "1,000,000 degrees outside," the listener may infer that it is "very hot" outside, as opposed to it actually being "1,000,000 degrees." Thus, the anchor would act as a more plausible exemplar (e.g., 125 degrees). Other anchors (e.g., 500,000 and 1,500,000 degrees) would produce similar effects, resulting in an asymptoting, rather than inverted-U, relationship between anchor extremity and the size of anchoring effects (Mussweiler & Strack, 2001). AIM can capture this asymptoting behavior through the parameter $\lambda_{\max}$. Another special case could occur when an individual has no prior information about a target item. If there is no prior representation, then the posterior representation will be made up entirely of the anchor information. In such a case, the posterior representation will be linearly related to the anchor.

For our purposes, we will use the flexible Bézier curve. The Bézier curve uses a set of control points to influence the path a curve may take. It requires the specification of at least two endpoints that the curve must pass through, and then any intermediate points can influence the shape of the curve, but the curve will not pass through these points – except in trivial cases. These requirements turn out to be helpful in our case, because we can directly build in restrictions on λ and ψ in the specification of the control points.

While there are many forms of Bézier curves, we focus on the set of rational Bézier curves that add adjustable weights to a set of polynomials. The weights in the specification of the Bézier curve will serve as the parameters in the mapping function. The mapping function will relate the prior consistency variable ψ (forming the x-axis) to $B(\psi)$, the proportion of $\lambda_{\max}$ (forming the y-axis), so that $\lambda = B(\psi)\lambda_{\max}$. Thus, the mapping function is defined on the unit square that ranges from zero to one on both the x and y axes.

We begin by specifying both the number $n+1$ and locations P of the control points. As the number of control points increases, the Bézier curve becomes increasingly more flexible, but the contribution of the individual points becomes more difficult to interpret. Here, we will use only four control points ($n = 3$) at the vertices of the unit square so that

$$P = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix},$$

where the ith row ($i = \{0, 1, 2, 3\}$) of P (denoted P_i) represents the ith (x, y) pair. The order of the control points affects the shape and orientation of the curve. For our mapping function, we require that the curve passes through the points $P_0 = (0, 1)$ and $P_3 = (1, 1)$, so that the curve transitions downward toward the point (0.5,0) depending on weights given to the intermediate control points P_1 and P_2.

Each control point P_i must also have an associated weight w_i. We set the endpoints w_0 and w_3 equal to one and set the intermediate points w_1 and w_2 equal to ξ so that $w = \{1, \xi, \xi, 1\}$. Constraining $w_1 = w_2$ does suggest that participants are unbiased with respect to the direction an anchor is presented. That is, anchors presented below the prior representation such that $\psi = \psi_1$ have just as much influence on the posterior representation as the presentation of an anchor that is above the prior representation such that $\psi = 1 - \psi_1$. We set $w_1 = w_2 = \xi$, although we acknowledge that this may be an overly restrictive assumption.[8]

Once the set of control points and associated weights have been specified, the mapping function is calculated by the equation

$$B(\psi|P, w) = \frac{\sum_{i=0}^{n} b_{i,n}(\psi) P_i w_i}{\sum_{i=0}^{n} b_{i,n}(\psi) w_i}, \tag{3}$$

where $b_{i,n}(\psi)$ denotes a set of polynomials, known as the Bernstein basis polynomials, which are given by

$$b_{i,n}(\psi) = \binom{n}{i} \psi^i (1-\psi)^{n-i}.$$

Fig. 9 illustrates the effects of anchor consistency on the adjustment from the prior representation to the posterior representation for five different values of ξ: 0, 0.2, 1, 2, and 30. The left panel shows the mapping function $B(\psi|P, w)$ (y-axis) for a range of ψ values (x-axis). The relationship between the prior consistency variable ψ and the anchor influence parameter λ is given by the equation

$$\lambda = \lambda_{\max} B(\psi|P, w). \tag{4}$$

[8] Although we believe that high and low anchors have comparable influences, we acknowledge that there are often additional constraints (e.g., zero quantities) imposed on low anchors.

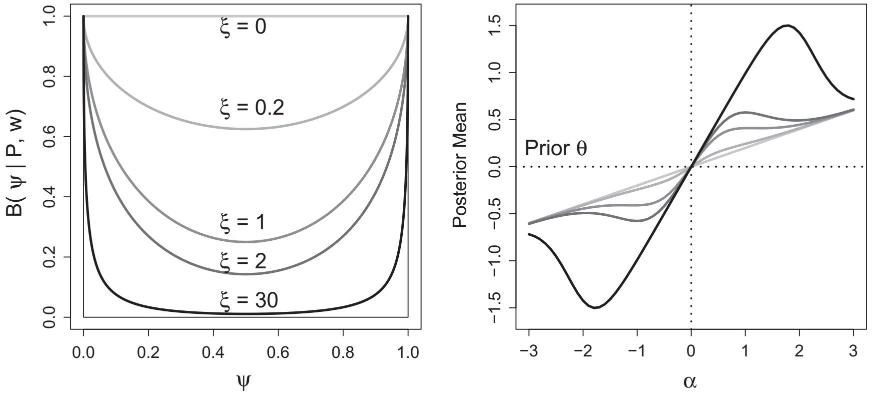

Fig. 9. The effects of consistency on anchor adjustments. The left panel shows the mapping function relating the prior consistency variable ψ to the anchor influence parameter λ at five different values for ξ. The right panel shows the effects that anchor consistency has on the adjustment from prior ($\theta = 0$ and $\delta = 1$) to posterior (y-axis) against a range of anchor values (x-axis).

As ψ increases from 0 to 0.5, the value of λ becomes small, and as ψ increases from 0.5 to 1, the value of λ increases again. Fig. 9 shows that as ξ increases from 0 to 30, λ decreases sharply, resulting in a stronger dependency on the information provided by the anchor. Thus, the influence of ξ becomes essential for understanding how consistency might play a role in a participant's judgment following the presentation of an anchor. To illustrate this effect, the right panel of Fig. 9 shows the degree of posterior adjustment made as a function of the five values of ξ in corresponding colors. To calculate the posterior adjustment, we assumed the prior mean $\theta = 0$ (represented as the dashed horizontal and vertical lines), the prior uncertainty $\delta = 1$, and $\lambda_{max} = 2$ (for illustrative purposes). Given these inputs and the level of ξ, we used Eqs. (4) and (B.11) to determine the mean of the posterior representation (y-axis) for a range of hypothetical anchor values α (x-axis). Fig. 9 shows an interesting U-shaped pattern that is modulated by the levels of ξ. Starting near the mean of the prior representation $\theta = 0$, we see little adjustment because the anchor value does not provide information that is unique to the prior. Hence, the mean of the posterior representation is similar to the mean of the prior representation. However, as the anchor moves toward the tails of the prior representation, we see that the influence of the anchor is more pronounced – especially for large values of ξ – moving the posterior representation away from the prior representation. However, at some point, the influence of the anchor declines and AIM reverts back to forming a posterior representation that is more consistent with the prior representation.

5.4. Summary

In short, we apply two constraints to extend our within-participant version of AIM First, we assume that confidence is related to the prior uncertainty parameter. For example, in Experiment 2 we will assume that across a single continuum Θ, confidence judgments behave similarly across individuals. This allows the model to infer how confidence may relate to uncertainty across both items and participants. Second, we use a measure of prior consistency to relate the influence of an anchor to a participant's prior information. In sum, as an anchor becomes more inconsistent with a participant's information (i.e., becomes more implausible), our AIM assumes that the anchor will have a proportionally smaller influence on that participant's final judgment about a target item.

Fig. 8 illustrates how the model works for three hypothetical anchors. The top left panel shows the prior representation with $\theta = 0$ and $\delta = 1$ along with the three anchors $\alpha = \{-3, -0.5, -1\}$ represented as vertical lines. After the presentation of these anchors, the prior consistency is calculated and this value is then converted to the anchor influence parameter. The top right panel shows the conversion of the three anchor values to $\psi = \{0.001, 0.309, 0.841\}$ (x-axis) along with the corresponding $B(\psi|P, w) = \{0.996, 0.360, 0.600\}$ (y-axis) where $\xi = 1$ and $\lambda_{\max} = 1$. This particular value of $\lambda_{\max}$ is unrealistically small, and was chosen for illustrative purposes. For example, the anchor A1 is very inconsistent with the prior and should have little influence on the posterior representation. However the bottom panel of Fig. 8 shows that the posterior does differ substantially from the prior. Despite this, the mapping function is meant to be a *relative* translation of the prior consistency, and so while the posterior does differ from the prior in this example, the difference observed for anchor A1 is small relative to anchors A2 and A3. The values of $B(\psi|P, w)$ are then used to compute λ by Eq. (4), which then creates the information functions corresponding to each of the anchor values shown in the middle panel. Finally, the information functions are combined with the prior representations to form the posterior representations, which are shown in the bottom panel.

6. Experiment 2

The primary goal of Experiment 2 was to test AIM's ability to account for empirical data on an individual participant basis. Experiment 2 employed a within-participant anchoring design to investigate how individuals integrate prior information and information provided by the anchor in the formation of a posterior judgment. Employing a within-participant design will provide the conditions necessary to fully identify the model we presented above.

6.1. Methods

Participants were recruited from Amazon's Mechanical Turk website and completed the task in exchange for \$0.50. The sample ($N = 189$) was 52% female with a mean age of 35 years. Participants completed a similar task to the regular anchors condition of Experiment 1. For each of the six questions from Experiment 1, the task began by eliciting an unanchored (i.e., prior) judgment about the target item and corresponding confidence rating for this first judgment, from 1 (*Very Unconfident*) to 13 (*Very Confident*). After completing this first task, participants completed roughly 15 min worth of unrelated non-numeric (to avoid the possibility of incidental anchors) tasks aimed at decreasing the accessibility of participants' unanchored judgments within their memories. Subsequently, participants were presented the anchor in a comparison question (e.g., is the length of the Mississippi River greater than or less than 3500 miles?) before estimating the target and providing a confidence judgment the second time for each of the six items.

6.2. The model

In contrast to Experiment 1, in Experiment 2 participants elicited an estimate from their prior representation and their posterior representation for each of the six questions. Thus, we are able to estimate parameters in the model associated with the prior representation (θ and δ) as well as the posterior representation (λ). In addition, we are now able to use a model that incorporates individual differences among all three parameters of AIM θ, δ, and λ. To do so, we will use a hierarchical version of AIM which will again be fit to the data in the Bayesian framework. The technical details of the hierarchical Bayesian model are presented in Appendix B, and we have provided JAGS (Plummer, 2003) code for fitting AIM to data in Appendix D.

Fig. 10 shows a graphical diagram for this model. To facilitate a comparison of this extended model to the basic model used for Experiment 1 (see Fig. 2), we added light gray circles around the nodes that are consistent with the previous model. As an extension of the earlier version of our model, we have added constraints from the confidence judgments as well as employed the prior consistency variable ψ to identify the model and dissociate the separate roles of prior uncertainty and the influence of the anchor values. Notice that for this model, in addition to a plate for each of the questions, we also have a plate for each of the participants and a plate for each of the confidence categories. At the first layer of the model, we have all of the participant parameters, namely, any of the parameters located on the "participant" plate. For example, the node for ξ_j is only on the participant plate, making it intrinsic

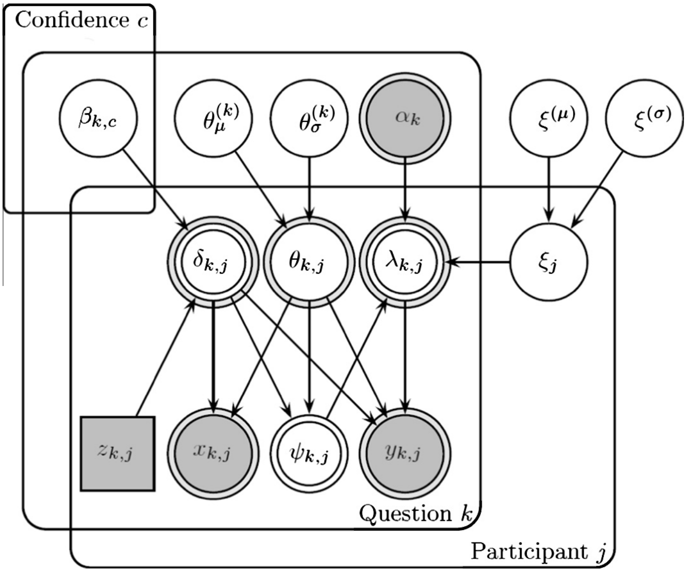

Fig. 10. A graphical diagram of the within-participant version of AIM. Nodes with explicit dependencies are connected with arrows, and plates are used to indicate a loop over an index. Nodes corresponding to fixed quantities, such as data or the anchor value, are shaded light gray, whereas unknown parameters are not shaded. Continuous variables are represented as circular nodes whereas discrete variables are represented as square nodes. To facilitate a comparison of this extended model to the basic model used for Experiment 1 (see Fig. 2), we added light gray circles around the nodes that are consistent with the previous version of the model.

to each participant but fixed across all questions. Similarly, the variable $\beta_{k,c}$ changes across each confidence category *and* anchoring question, but is held fixed across all participants.

6.3. Results

We fit the model in a Bayesian framework by again using DE-MCMC sampling (see ter Braak, 2006; Turner et al., 2013, for details). We ran a blocked version of the algorithm with 40 chains distributed across 8 processors for 3000 iterations after a burnin period of 4000 iterations, resulting in 120,000 total samples from the joint posterior distribution. Standard procedures were used to ensure convergence.

There are a number of parameter analyses and posterior checks we could present here; however, for brevity we will instead focus on two sets of parameters that are central to the development of our model: the prior uncertainty parameters β and the anchor sensitivity parameters ξ. Similar to the analysis in Experiment 1, we will then examine the posterior predictive distributions of the model against the data that were observed.

6.3.1. Confidence and prior uncertainty

We begin by assessing the relationship our model parameters have to the degree of confidence reported for each individual judgment. In the model section, we hypothesized that as the reported confidence increased, we should see a much smaller value for the prior uncertainty parameter. Recall that the model for these data assumed a separate prior uncertainty parameter for each question by confidence judgment, and assumed no relationship between the confidence parameters. However, to apply constraints to the model at the individual level, we assumed that participants in the kth question who reported a confidence judgment of c shared the same prior uncertainty parameter $\beta_{k,c}$, but may not necessarily share the same prior mean $\theta_{k,j}$.

Fig. 11 shows the estimated posterior distribution for each parameter $\beta_{k,c}$ on the log scale. The separate panels represent the estimates for each of the six questions (i.e., the rows of the matrix β), and the separate violin plots within a panel represent the elicited confidence judgment (i.e., the columns of the matrix β). There is a strong negative relationship between prior uncertainty and the confidence response. Specifically, Fig. 11 shows that as the reported confidence judgment increases from 1 to 6, the degree of uncertainty in the prior representations for this group of participants decreases, reflecting a more peaked prior representation (see Fig. 1). In general, we find that as confidence increases, the difference between the prior judgment and the posterior judgment decreases, and our estimates of $\beta_{k,c}$ reflect this relationship.

One compelling advantage of the Bayesian approach is the quantification of uncertainty in each parameter estimate. For example, the top left panel of Fig. 11 shows that the estimate for $\beta_{1,6}$ (the violin plot furthest to the right) is much more variable than the estimate for say, $\beta_{1,1}$ (the violin plot furthest to the left). Thus, while we can be sure that the true value for $\log(\beta_{1,1})$ is somewhere between -1 and 1, the true value for $\log(\beta_{1,6})$ might range anywhere from -1 to 3. The uncertainty here is largely due to the sparsity in the number of confidence judgments of "6". However, other factors, such as the variability in the estimate itself, will influence the shape of these posterior distributions.

6.3.2. Raw mean adjustment and model parameters

The next parameter we investigate is the anchor sensitivity parameter ξ. The parameter ξ was constructed to provide some insight into how an individual might update for any hypothetical anchor by relating the measure of prior consistency (see Eq. (2)) to the spread of the information function. The parameter ξ controls this mapping so that as ξ increases, anchors are perceived to carry more information than when ξ is small (see Fig. 9). Thus, participants whose posterior judgments differed considerably from their prior judgments should have large values of ξ.

To construct a holistic measure of how much a participant's prior and posterior judgments differ, we first calculate the average absolute difference (i.e., their average updating score), given by

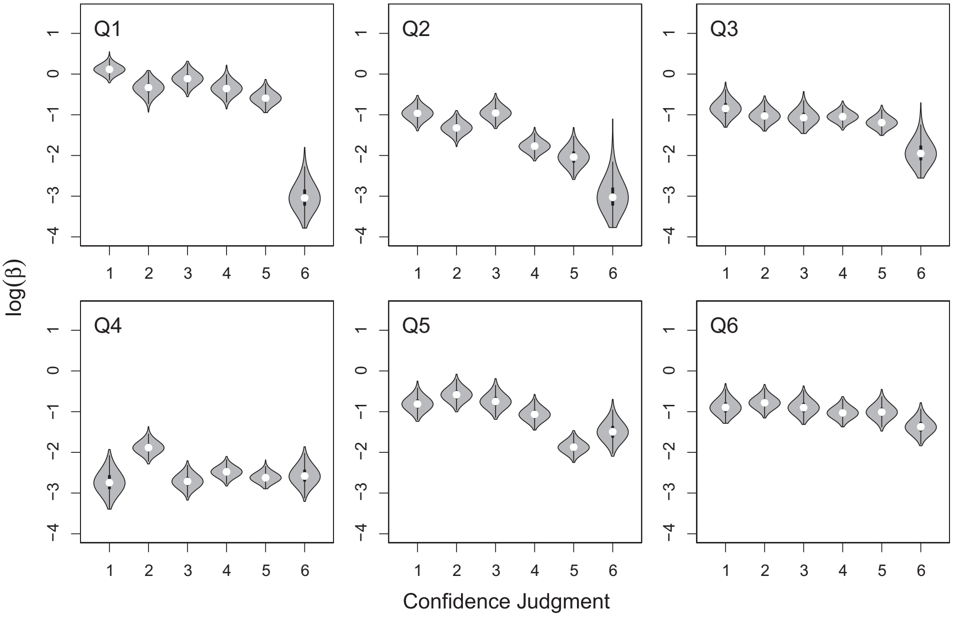

Fig. 11. The estimated posterior distributions for β (y-axes) on the log scale for each of the six questions as a function of confidence response (x-axes).

$$MA_j = \frac{1}{6}\sum_{k=1}^{6}|x_{k,j} - y_{k,j}|.$$

Although this measure will not provide us with insight about *why* a participant might update their estimates, it provides a measure that we can compare ξ against. The left panel of Fig. 12 shows the average absolute differences *MA* against the posterior estimates for ξ for each participant. Each ξ_j is shown as the median, represented as the black dot, and the 95% credible set, represented as the length of the vertical lines. The left panel shows a positive relationship between *MA* and ξ: as ξ increases, the average absolute differences also tends to increase. The variable *MA* is correlated with the parameter ξ, resulting in a correlation of $r = 0.27$ ($N = 166$, and $p = 0.001$). This result indicates that ξ is related to the degree that individuals' update their beliefs when presented an anchor (i.e., their susceptibility to the anchor). The right panel of Fig. 12 shows *MA* against the posterior estimate for β, collapsed across questions. Each node represents the confidence category selected, and is labeled accordingly. The general pattern, also reflected in Fig. 11 is that as β increases, the degree of adjustment from prior to posterior representations increases ($r(6) = 0.86, p = 0.028$). In other words, the more certain people were a priori, the larger were their adjustments.

While the variable *MA* does provide a reasonable statistic for the average degree of movement for a participant, it does not provide any information about the reason for the updating. AIM, on the other hand, has two mechanisms to explain this updating. The first rests in the degree of prior uncertainty, which we examined above. The second is anchor susceptibility, which we have just discussed. We have shown that together these estimates provide a reasonable interpretation of why participants update their prior judgments after the presentation of an anchor. In particular, the range of the estimates for ξ combined with the relationship of prior uncertainty and elicited confidence suggest that both prior uncertainty and anchor susceptibility have significant effects on the updating of the prior representation to the posterior representation.

6.3.3. Predicting individual anchoring effects

We can also make predictions for how much a participant will adjust their posterior representations given a hypothetical anchor value and their pattern of behavioral data in Experiment 2. In this within-participant version of AIM, each subject has an anchor susceptibility parameter ξ that tells how impactful an anchor will be given its location within the prior representation. This mapping function allows us to generate predictions for new questions, assuming we know what the prior representation looks like (e.g., see Experiment 1). Suppose we have a new item that, from previous data collection we have learned is centered at $\theta = 0$. The prior uncertainty of this item varies according

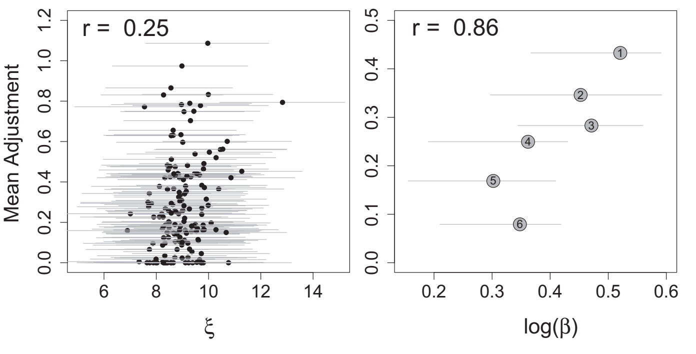

Fig. 12. Relationships between the mean adjustment from prior to posterior representations against the model parameters. The left panel shows the mean adjustment from prior to posterior against the estimated posterior distribution for ξ across all questions for each subject. The right panel shows the mean adjustment from prior to posterior averaged across questions against the posterior distribution of each confidence response β. The median of the estimated posterior is shown as the black dot and the 95% credible set is shown as the height of the vertical gray line. In the right panel, the reported confidence value corresponding to each β is labeled on each dot.

to the degree of confidence reported by a participant, in an analogous manner as that of Fig. 11, where higher degrees of confidence translate into smaller values for δ (i.e., greater prior certainty). Using these two sources of information, we can now generate predictions for anchoring effects on novel items.

Fig. 13 shows the predicted anchoring effects for a range of hypothetical anchor values for four representative subjects. These subjects were chosen as a representative sample, where the Subject 22 was the least susceptible participant in our data ($\xi = 6.90$), Subject 150 was the most susceptible participant in our data ($\xi = 12.83$), and Subjects 94 and 149 spanned values of ξ in between ($\xi = 9.06$ and $\xi = 9.99$, respectively). In each panel, the degree of posterior adjustment is shown as a function of three levels of reported confidence, color coded according to the legend in the top left panel. As in Fig. 9, the mean of the prior representation $\theta = 0$ is illustrated as the dashed lines. To select the δ parameters for this new item, we simply used the estimates of β we obtained from Question 6 of Experiment 2 (i.e., see Fig. 11). Fig. 13 shows that AIM predicts that with less prior certainty, the more susceptible participants are to the influence of the anchor, especially when the anchor is presented in the range of (0.5, 1.5). Across panels, we see a very intuitive pattern of posterior adjustments across the range ξ, namely that as ξ grows, the degree of posterior adjustment also grows. Together these predictive results suggest that both the degree of prior certainty and the level of individual-specific anchor susceptibility influence the magnitude of adjustment from prior to posterior representations.

6.3.4. Predicting group anchoring effects

As a final assessment of model fit, we compared the posterior predictive distribution of our model to the data that were observed for each anchoring question. Fig. 14 shows these posterior predictive distributions (gray clouds) along with the data (black dots) for each question (represented as separate panels). For each question, we first plotted the prior judgment $x_{k,j}$ against $y_{k,j}$ observed in the data. After fitting the model to these data, we then generated 500 samples from the posterior predictive dis-

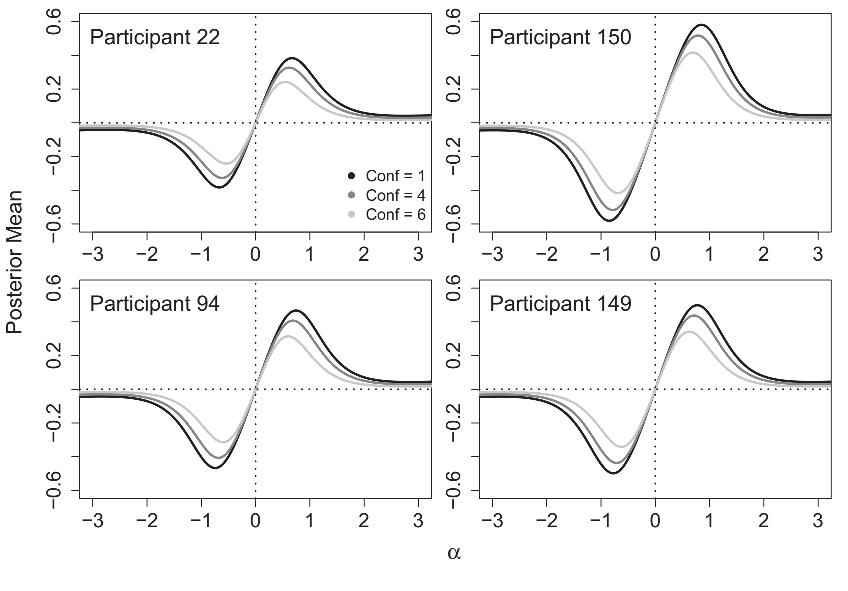

Fig. 13. Predicted anchoring effects for four representative participants in Experiment 2. Each panel corresponds to an individual participant, and predictions are shown for three different confidence levels, color coded according to the legend in the top left panel. In each panel, the y-axis shows the predicted mean adjustment from prior to posterior representation against a continuum of possible anchor values (x-axis). The mean of the prior representations are illustrated with the dashed lines (i.e., $\theta = 0$).

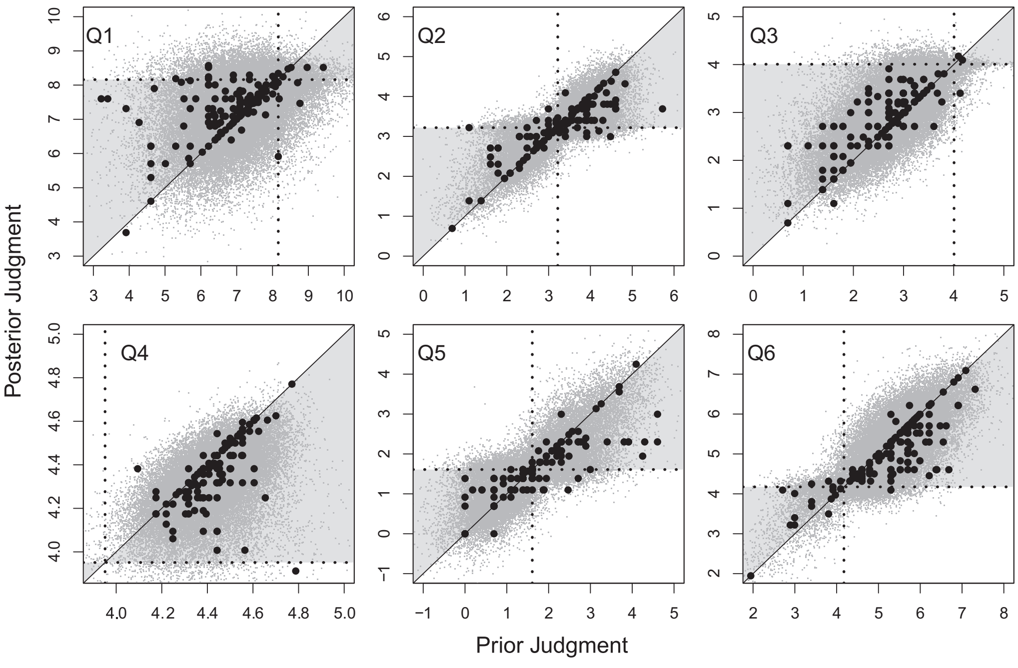

Fig. 14. The posterior predictive distributions (gray cloud) along with the observed responses (black dots) for the prior judgment (x-axes) and posterior judgment (y-axes) for each of the six anchoring questions in Experiment 2. Both axes are log-transformed. The dashed lines represent the anchor value that was presented. The light gray triangular regions indicate positive updating (i.e., updating from prior estimates toward the anchor value).

tribution.[9] Fig. 14 shows that AIM does an excellent job of matching both the spread and relative density of the observed data.

The vertical and horizontal dashed lines in each panel shows the anchor value that was presented. We also shaded the regions that correspond to an updating from the prior representation toward the presented anchor value. For example, in the top left panel of Fig. 14 the bottom left region is shaded to indicate that if a prior judgment of say, 5.1 increased toward the presented anchor value of 8.16 (the horizontal dashed line), it would lie somewhere in the interval (5.1,8.16). Similarly, the top right triangular area is shaded to reflect the situation where a participant's prior judgment is above the presented anchor. Following the presentation of 8.16 (i.e., the log-transformed anchor value), if the participant's posterior judgment shifted downward (and hence would be under the diagonal line $y = x$) toward the anchor value, their data point would fall in the gray triangular region. Simply put, these gray triangular regions reflect the predictions from extant qualitative theories (i.e., in the presence of an anchor, individuals will update their judgments in the direction of the anchor, and that the magnitude of the adjustment is unknown).

As in the individual-level predictions (see Fig. 13), Fig. 14 shows that the distribution of responses as a function of distance from the anchor is curvilinear: there tends to be small updates in the posterior estimates when the prior estimates are nearby or far from the anchor, but large updates (i.e., anchoring effects) in between these extremes. This curvilinear pattern is consistent with the inverted U-shaped relationship observed by Wegener et al. (2001).

The figure also shows that AIM can account for reversals, where the judgment from prior to posterior moves in the opposite direction of the anchor value or moves to the other side of the anchoring

[9] While 500 samples may seem like a small number, it was chosen for graphical purposes. Increasing the number of samples produced a dense cloud of points making it difficult to assess the density at any particular location.

value. Recall that in AIM, because the response is assumed to be sampled from a representation (either the prior or posterior), there is some probability of a reverse anchoring effect, where the judgment moves in the opposite direction of the anchor. Although this pattern of effects may seem unlikely – indeed, they are inconsistent with current qualitative theories of anchoring – such outcomes are evident in each of the questions in Fig. 14 (i.e., when a black dot falls outside of the shaded regions). In addition, AIM naturally captures the effect that the closer a prior judgment is to a presented anchor value, the smaller the difference between the prior and posterior judgments. For example, the top middle panel shows that in areas far from the anchor value, the variability in the joint distribution of the prior and posterior representations is higher than in regions near the presented anchor (i.e., regions near the intersection of the vertical and horizontal dashed lines). This pattern is also exhibited in the behavioral data.

6.4. Summary

Experiment 2 was intended to test the adequacy of the within-participant version of AIM, assessing the extent to which individuals integrate anchor information into their prior representation to form a posterior representation of the target item. Although it is not feasible to actually measure both an individual's prior and posterior representation, Experiment 2 provides a novel insight into this information integration. To the best of our knowledge, this is the first demonstration of the standard anchoring effect at the individual level. Although participants' posterior estimates were heavily biased by their prior estimates, the current results do show anchoring effects at the individual level. Furthermore, the within-participant version of AIM provided impressive fits to the data at the individual level.

7. General discussion

Four decades of anchoring research has demonstrated the stability of anchoring effects (Mussweiler, 2001), and the important role of this class of effects on everyday judgments and decisions (for reviews, see Epley, 2004; Furnham & Boo, 2011). The current article presented two versions of AIM, a descriptive model for quantifying anchoring effects. Given that AIM is intended to be a descriptive model of anchoring effects, one may wonder, when is AIM an appropriate tool for my data? What specific steps should I take to implement AIM? In the next section, we provide some conceptual guidelines for using AIM, as well as a discussion of what benefits it provides.

7.1. Using AIM: a conceptual overview

First and foremost, AIM is intended to provide a simple characterization of the degree to which a piece of information affects a quantitative judgment. Although there are many judgment contexts in which the model would be appropriate, the current article is focused specifically on anchoring effects. AIM is agnostic regarding the type of target (e.g., lengths of rivers, distances between cities, or even something more abstract such as the credibility of a witness testimony). The target is not limited to estimates (e.g., "What is the length of the Mississippi River?") and can include observable behaviors (e.g., amount consumers repay toward credit card debts; Navarro-Martinez et al., 2011; Stewart, 2009). AIM is intended to fit data for targets used in the previous literature as well as a range of other targets as well.

7.1.1. The between-participant version of AIM

For the between-participant version of AIM to be identifiable, AIM requires that some judgments are provided in a "prior" condition, and at least one "anchored" condition. The prior condition involves collecting unanchored judgments of the target item. For example, data for the prior condition could include asking participants to estimate the length of the Mississippi River or observing consumers' credit card payments. Such data allow AIM to estimate θ, the prior mean parameter, and δ, the prior uncertainty parameter. In the same manner as Experiment 1, these parameters can be inferred by assuming that the mean and the variance of the observed distribution of data in the prior condition represent a reasonable proxy for the average mean and variance of an individual's prior representation in the anchored condition.

The anchored condition involves collecting data about the target item after the presentation of an anchor. For example, asking participants to estimate the length of the Mississippi River after asking them to judge whether the Mississippi River is less than or greater than some anchor value. AIM is agnostic to the type of anchor and only requires its user to be aware of the approximate distribution of information provided by the anchor. For example, if an acquaintance tells you that the Mississippi River is less than 3500 miles long, AIM would require the use of the truncated information function from Section 3.2 of this article.[10] Any piece of quantitative information presented that may influence participants' judgments of the target item can be considered an anchor within AIM. Using posterior judgments (i.e., judgments after the presentation of the anchor), AIM can estimate the anchor influence parameter λ.

Using the both the data we collect from the prior and anchor conditions and the code provided in Appendix C, we can obtain estimates for each of the parameters: θ, δ, and λ. As we discussed, the data from the prior condition constrains all three parameters, whereas the data from the anchored condition constrains only λ.

Presumably, the procedure for using the between-participant version of AIM could also be applied to within-participant experimental designs as well. Although the within-participant version of AIM allows for improved fits (i.e., the ability to model individuals rather than groups), when there is not enough data to sufficiently constrain the within-participant version of AIM, one could apply the between-participant version to the data. For example, imagine observing consumers' credit card payments at time t_0; these data could be treated as the prior condition. On the following credit card statement, these individuals are offered "recommended payment amounts." If we observed the payments at time t_1, these data could be treated as the anchored condition. Therefore, the between-participant version of AIM could still be applied to a within-participant design, taking care to acknowledge obvious dependencies in the data.

7.1.2. The within-participant version of AIM

As we have shown in Experiment 2, AIM can also be used to quantify anchoring effects on an individual level. Although we have not yet conducted extensive model identifiability analyses, we can provide some recommendations about how to constrain AIM appropriately. First, we require data from two or more anchoring conditions along with the prior condition. These extra conditions worth of data help to identify the parameters of the linking function (i.e., the parameters of the Bézier curve) used by AIM to articulate the updating process from prior to posterior representations. Second, some extra source of data is needed to constrain the parameters that correspond to the prior representation. In the within-participant experimental design, we have only one observation per subject from the prior condition. Given that AIM quantifies the prior representation via two parameters, one other source of constraint is needed. Although hierarchical constraints across participants is one straightforward solution, we recommended obtaining confidence measures in the prior condition. This could be done directly by assessing participants' subjective confidence, or indirectly by measuring or observing data that can serve as a reasonable proxy for participants' prior uncertainty. Our results from Experiment 2 suggest that participants are able to provide explicit information about how uncertain they are in the judgment they provide, and this uncertainty manifests in the amount of updating from prior to posterior judgments. Hence, the confidence measures provide a simple mechanism that constrains AIM well enough to extract similar types of information that is obtained in the between-participant version of AIM.

7.2. Limitations

A major limitation of the current research is that it is difficult to obtain enough data to appropriately identify AIM. For the between-participant version of AIM, we assessed the unanchored judgments of a

[10] We do not argue whether or not a friend's recommendation would be considered an anchor within the existing literature, only that such recommendations could be modeled within AIM.

control group and used the distribution of responses as an approximation for participants' average prior representations in the experimental conditions of Experiment 1. Experiment 2 attempted to remove this assumption by collecting evidence for both a participant's prior and posterior representation of each target judgment. However, once an individual is asked to estimate a target item, his or her estimate of the target item will act as a self-generated anchor, making it difficult to differentiate between the effects of the self-generated anchor and the anchor value presented in the subsequent question. Furthermore, it seems unreasonable to expect that a participant will respond differently to the second presentation of a question when it is essentially the same as the first presentation. Regardless, we attempted a within-participant manipulation in Experiment 2, in which each participant was asked to report a value from their prior representation, and then report an estimate from their posterior representation, after the presentation of an anchor. As predicted, some participants did respond similarly (some even provided the same response) for each question. After fitting AIM to the data, we found that our model could account for these data, despite the large dependencies in participants' responses. AIM predicted a degree of variability that was consistent with each participant's responses. Although we are skeptical about the effectiveness of within-participant manipulations, our data exhibited moderate anchoring effects across participants, and AIM was successful in accounting for these data.

One limitation of AIM is the restriction of the form of both the prior representation and the anchor information function. We have already expressed concerns about these assumptions, but we reemphasize that more complicated priors may be necessary in certain instances (e.g., Lewandowsky et al., 2009). When this is the case, one could employ a variety of Bayesian techniques such as finite or infinite mixture priors, or empirical priors such as a kernel density estimate of the data from the prior representation (see Silverman, 1986).

In the within-participant version of AIM, one potential limitation lies in the mapping function. In this article, we proposed a mapping function that was very simple, symmetric, and provided a convenient interpretation for the parameter ξ. However, we suspect that this may be one component of the model that could be improved. Furthermore, the mapping function's dependence on the upper bound parameter $\lambda_{\max}$ also poses a slight complication. In this article, we simply proposed that $\lambda_{\max}$ be estimated prior to fitting AIM to the data. Although we have not tested this directly, one could incorporate $\lambda_{\max}$ into the model and estimate it.

A final potential limitation of AIM is in the way it integrates information in the prior representation and the anchor to form a posterior representation. In the current article, we have only considered a Bayesian form of updating these two sources of information, but it may well be the case that simpler forms of integration provide a better account of the data (e.g., Lieder et al., 2012). For example, one can imagine a simple weighted average of the mean parameter θ with the anchor value α to form the posterior representation with some error associated with the decision process. In this framework, the weight applied to the anchor value would resemble the parameter λ, and the weight assigned to the prior representation would resemble the parameter δ. Weighted averages certainly require less of a computational burden on the observer compared to full integration, but by virtue of their simplicity weighted average models may lack the required flexibility to fit empirical data. Rigorous tests might reveal the relative strengths and weaknesses of various integration strategies, but such tests are one step beyond the initial goals of the present article, and so we will save these investigations for future research.

7.3. Future directions

We have demonstrated that AIM can account for a variety of anchoring effects. Due to the information integration proposed in this article, we would also expect that AIM could be extended to account for dynamic judgments as well. Dynamic judgments are those in which individuals incorporate several sources of information and make a series of estimates. For example, an extension of AIM could account for and reliably predict the learning processes in stock trading, gambling, and related dynamic environments.

Although we consider AIM to be a step forward in modeling anchoring effects, we hope that AIM spurs additional research into the development of methods that more appropriately collect within-participant data, and appropriate models for fitting the data. In addition to traditional methods for studying anchoring effects, we believe that future anchoring research should concern within-participant data. Such data would allow for greater insight into the psychological processes underlying anchoring effects.

7.3.1. AIM as a predictive model for "Nudging"

Although forty years of research have already been dedicated to the anchoring phenomenon, the current article opens up several novel avenues for additional research. One of the most substantial opportunities involves extending AIM from a descriptive model into a predictive model. In its current form AIM provides a useful tool for quantifying the observed relationships present in the data. We use the data from the current experiments to estimate the anchor-influence parameter λ. If λ were known, then for any given prior, AIM could estimate predictions for the posterior representation corresponding to any presented anchor. For instance, if prior and posterior data were collected across a range of anchor values within a given paradigm, AIM could infer λ corresponding to any anchor value and predict, a priori, the posterior representation for any anchor value. As an example, past research demonstrated that the listing price of a home acts as an anchor on subsequent home valuations (Northcraft & Neale, 1987). For a given market, incorporating existing priors (e.g., past listing prices) and posteriors (i.e., corresponding offers), AIM could predict a posterior for any given anchor. We find that the parsimony of AIM allows it to be tractable in a myriad of domains.

The ability to predict, not just the existence of anchoring effects, but a priori estimates of the magnitude of anchoring effects would provide a valuable tool. Considerable research on "nudging" (also referred to as choice architecture) has highlighted the utility of behavioral insights for public policy (Allcott & Mullainathan, 2010; Johnson & Goldstein, 2003; Thaler & Sunstein, 2008). By using behavioral interventions, researchers and policy developers can construct decision environments that facilitate effective judgments and decisions. Given the reliability and prevalence of anchoring effects across a range of important domains, AIM may provide an important tool for policy. In its current form, the anchoring literature can inform policy about whether or not an anchoring effect will occur in a particular paradigm. Furthermore, AIM offers the ability for policy developers to potentially choose "optimal" anchors (Urminsky & Goswami, 2015).

7.3.2. AIM as a tool for studying multiple anchors

Despite the robustness of the anchoring phenomenon, very little research has focused on how multiple anchors influence judgments (Whyte & Sebenius, 1997; Zhang, Li, & Zhu, 2011). Given that many judgments and decisions involve more than one anchor (e.g., minimum payment and suggested payment on credit card statements), understanding the role of multiple anchors seems an important endeavor. When using traditional approaches, identifying the unique influences of each anchor in a multiple-anchor paradigm is unfeasible (i.e., two inputs but only one observation). Utilizing the hierarchical structure within AIM, a version of our model could be adapted to model the influences of multiple anchors on the posterior representation.

7.3.3. AIM as a mechanism for differentiating anchoring theories

The anchoring-effects literature is composed of five non-mutually exclusive theories, each of which provide useful insight toward understanding the processes underlying anchoring effects. We do not discount the existing literature, but believe that the literature needs trimming to improve its usability, both for theorists and practitioners. Because existing anchoring theories differ in terms of how anchors provide information (i.e., the second component of anchoring effects discussed in the introduction), we suggest that theorists could build process-based models of their proposed psychological mechanisms within the AIM framework and test for the model's adequacy in accounting for the data. For example, Bhatia and Chaudhry (2013) introduced an associative network model of how exemplars are retrieved in a manner consistent with the selective accessibility hypothesis (Mussweiler & Strack, 1999; Strack & Mussweiler, 1997). A version of this model could be fit within AIM to determine the anchor-influence parameter λ. In doing so, AIM would become a process-based model that allows for testing the adequacy of the selective accessibility hypothesis. Because AIM – as a descriptive model – is agnostic about how the anchor provides information, AIM would be a valuable tool for discriminating the efficacy of differing accounts. At the very least, this approach may not rule out any of the psychological mechanisms proposed within the literature, but could provide evidence as to the prevalence of any of the proposed psychological processes across a range of paradigms within the anchoring literature. We do not discount the existing qualitative approaches for conceptualizing anchoring effects, rather we view AIM as a useful tool for further developing existing accounts. Thus, although AIM is descriptive in its current form, AIM provides a novel means for testing and differentiating theories.

8. Conclusions

The abundance of anchoring research highlights its prevalence and stability as a core feature of human judgment and decision making. Anchoring effects often involve a quantitative anchor and a quantitative estimate, as such qualitative theories of anchoring effects are insufficient for fully articulating this class of effects. The current article introduces AIM for the quantitative assessment of anchoring effects. AIM not only provides a means for assessing the magnitude of anchoring effects, but the opportunity for theory evaluation and prediction.

In this section, we present the equations necessary to fit the two versions of AIM to data. Our exposition is presented deductively, moving from more general equations which are applicable to any experimental data to the specific ones used to fit our data from Experiment 1 and Experiment 2, respectively. All notation, except where noted, is as presented in the model sections of the article.

Appendix A. The between-participant version of AIM for Experiment 1

To appropriately constrain the model, we assumed that responses elicited in the prior condition were reflective of the general population prior for each question. In this particular version of AIM, we are not concerned with how anchors affect *individual* behavior. Instead, we are only concerned with the effects particular anchors have on *groups* of individuals. In Experiment 2, we extend the basic model to incorporate individual differences.

A.1. The prior representation

We first applied a log transformation to the data for each question, and fit each question independently. The log transformation was useful in reducing skew for some of the items, and as a consequence, it helped to justify our normality assumption. We denote data from the prior condition for the jth subject as x_j. We assume that the J participants have the same prior representation from which they independently sample their estimate. Hence, the data x_j arise from a normal prior representation with mean θ and standard deviation δ, such that

$$\begin{aligned} x_j &\sim \text{Prior}(\theta, \delta) \\ &\sim \mathcal{N}(\theta, \delta), \end{aligned} \tag{A.1}$$

where $\mathcal{N}(a, b)$ denotes a normal distribution with mean a and standard deviation of b. To obtain the likelihood equation for a given combination of (θ, δ), we evaluate

$$f(\theta, \delta | x) = \prod_{j=1}^{J} \text{Prior}(x_j | \theta, \delta). \tag{A.2}$$

For this application, Eq. (A.2) is determined by passing the values x_j through the normal density function, but of course other prior representations are possible.

A.2. The information function

We then assume that, upon presentation of the anchor value α_k in the kth condition, the prior representation is updated through multiplication of the information function, which is governed by the anchor influence parameter λ_k. Here, we assume that the information function is normal in form, such that

$$\text{Info}(\Theta|\alpha_k, \lambda_k) = \frac{1}{\sqrt{2\pi}\lambda_k} \exp\left[-\frac{(\alpha_k - \Theta)^2}{2\lambda_k^2}\right]. \tag{A.3}$$

A.3. The posterior representation

Following the presentation of the anchor, the participant updates their prior representation by the information function to arrive at the posterior representation

$$\text{Post}(\Theta|\theta, \delta, \lambda_k, \alpha_k) = \frac{\text{Prior}(\Theta|\theta, \delta)\text{Info}(\Theta|\alpha_k, \lambda_k)}{\int \text{Prior}(\Theta|\theta, \delta)\text{Info}(\Theta|\alpha_k, \lambda_k)d\Theta}.$$

We let $y_{k,j}$ denote the jth response in the k anchoring condition, where $j \in \{1, \ldots, M_k\}$ and M_k is the number of responses obtained in the kth anchoring condition. For the posterior representation, we again assume that the data $y_{k,j}$ are independent and identically distributed. Hence, the data

$$y_{k,j} \sim \text{Post}(\theta, \delta, \lambda_k, \alpha_k).$$

Because the data $y_{k,j}$ are sampled from a posterior representation that depends on a latent prior representation, the parameters θ and δ controlling the prior representation are not fully observed. That is, the estimates for both θ and δ are obtained *jointly* from the data from the prior condition x and the data from the anchoring conditions y_k. Given the distributional assumptions above, the specific influence of the data from all participants in the kth condition, which we denote $y_{k,1:M_k}$, can be described by the equation

$$f(\theta, \delta, \lambda_k | y_{k,1:M_k}, \alpha_k) = \prod_{j=1}^{M_k} \text{Post}(y_{k,j}|\theta, \delta, \lambda_k, \alpha_k), \tag{A.4}$$

where α_k is assumed to be known. If we further assume that the data y are independent across the K conditions, then Eq. (A.4) becomes

$$f(\theta, \delta, \lambda | y, \alpha) = \prod_{k=1}^{K}\prod_{j=1}^{M_k} \text{Post}(y_{k,j}|\theta, \delta, \lambda_k, \alpha_k). \tag{A.5}$$

The assumption of normality in the prior representation and the normal information function enables us to provide closed form expressions for the posterior representation. By combining Eqs. (A.1) and (A.3), the posterior representation can be expressed as

$$\begin{aligned}
\text{Post}(\Theta|\theta, \delta, \lambda_k, \alpha_k) &= \frac{\exp\left[-\frac{(\Theta-\theta)^2}{2\delta^2}\right]\exp\left[-\frac{(\alpha_k-\Theta)^2}{2\lambda_k^2}\right]}{\int \exp\left[-\frac{(\Theta-\theta)^2}{2\delta^2}\right]\exp\left[-\frac{(\alpha_k-\Theta)^2}{2\lambda_k^2}\right]d\Theta} \\
&\propto \exp\left[-\frac{1}{2}\left(\frac{1}{\delta^2}+\frac{1}{\lambda_k^2}\right)\left(\Theta - \left(\frac{\frac{\theta}{\delta^2}+\frac{\alpha_k}{\lambda_k^2}}{\frac{1}{\delta^2}+\frac{1}{\lambda_k^2}}\right)\right)^2\right] \\
&= \frac{1}{\sqrt{2\pi}b_k}\exp\left[-\frac{(\Theta - a_k)^2}{2b_k^2}\right],
\end{aligned} \tag{A.6}$$

which is a normal distribution with mean a_k and standard deviation b_k, where

$$\begin{aligned}
a_k &= \frac{\frac{\theta}{\delta^2}+\frac{\alpha_k}{\lambda_k^2}}{\frac{1}{\delta^2}+\frac{1}{\lambda_k^2}}, \text{ and} \\
b_k &= \left(\frac{1}{\delta^2}+\frac{1}{\lambda_k^2}\right)^{-1/2}.
\end{aligned} \tag{A.7}$$

In other words, for this particular application, the data

$$y_{k,j} \sim \mathcal{N}(a_k, b_k).$$

A.4. Estimating the model parameters

The likelihood function for the parameters $\theta,\ \delta$, and λ, conditional on the anchor value α and the data x and y, is determined by combining Eqs. (A.2) and (A.5), such that

$$L(\theta, \delta, \lambda | x, y, \alpha) = \prod_{j=1}^{J} \text{Prior}(x_j | \theta, \delta) \prod_{k=1}^{K} \prod_{j=1}^{M_k} \text{Post}(y_{k,j} | \theta, \delta, \lambda_k, \alpha_k). \quad \text{(A.8)}$$

Because both the prior and information functions are normal in form, Eq. (A.8) becomes

$$L(\theta, \delta, \lambda | x, y, \alpha) = \prod_{j=1}^{J} \phi(x_j | \theta, \delta) \prod_{k=1}^{K} \prod_{j=1}^{M_k} \phi(y_{k,j} | a_k, b_k), \quad \text{(A.9)}$$

where a_k and b_k are evaluated using Eq. (A.7), and $\phi(x|a, b)$ denotes the normal density with mean a and standard deviation b, evaluated at the location x. With the likelihood function in hand, one can use standard routines to optimize Eq. (A.8) – or when appropriate, Eq. (A.9) – to obtain maximum likelihood estimates for the parameters $\theta,\ \delta$, and λ.

A.4.1. Bayesian inference

When estimating the parameters $\theta,\ \delta$ and λ in the Bayesian framework, one must provide prior distributions for each parameter. To fit the data from Experiment 1, we assumed a normal prior for θ and gamma priors for both δ and λ such that

$$\theta \sim \mathcal{N}(\mu, 10),$$
$$\delta \sim \Gamma(1, 1), \text{ and}$$
$$\lambda_k \sim \Gamma(1, 1).$$

The values for each hyperparameter μ were selected somewhat informatively for each item such that $\mu = \{7, 3, 3, 4.5, 1, 5\}$.

Once the priors for $\theta,\ \delta$ and λ have been specified, we can now write the joint posterior distribution as

$$\pi(\theta, \delta, \lambda | x, y, \alpha) \propto L(\theta, \delta, \lambda | x, y, \alpha) \prod_{k=1}^{K} \pi(\lambda_k) \pi(\theta) \pi(\delta), \quad \text{(A.10)}$$

where the likelihood is provided in Eq. (A.8). Because Eq. (A.10) is rarely in a convenient form, we generally must use algorithms to sample from it. In fitting the data from Experiment 1, we obtained samples directly from Eq. (A.10), but one could employ a Gibbs sampling algorithm by sampling from the conditional distributions for each of the parameters sequentially. A Gibbs sampling approach would be particularly advantageous when the dimensionality of λ is large, because of the decline in the acceptance rate of proposal generation in high-dimensional spaces (see Gelman, Carlin, Stern, & Rubin, 2004; Robert & Casella, 2004). The conditional distributions for $\theta,\ \delta$ and λ_k are

$$\begin{aligned}
\pi(\theta | \delta, \lambda, x, y, \alpha) &\propto L(\theta, \delta, \lambda | x, y, \alpha) \pi(\theta), \\
\pi(\delta | \theta, \lambda, x, y, \alpha) &\propto L(\theta, \delta, \lambda | x, y, \alpha) \pi(\delta), \text{ and} \\
\pi(\lambda_k | \theta, \delta, x, y, \alpha_k) &\propto L(\theta, \delta, \lambda_k | x, y, \alpha_k) \pi(\lambda_k) \\
&\propto \prod_{j=1}^{M_k} \text{Post}(y_{k,j} | \theta, \delta, \lambda_k, \alpha_k) \pi(\lambda_k),
\end{aligned}$$

respectively. In other words, the data x and y are explicitly dependent on the parameters θ and δ, but λ_k only affects the data y.

Appendix B. The within-participant version of AIM for Experiment 2

B.1. The prior representation

To simplify the notation, we will again divide the data up into data from the kth question from the prior question x_k (i.e., the presentation of Question k that does not contain an anchor) and the kth anchoring question y_k. To indicate individual responses, we add the subscript j to the appropriate elements. For example, $x_{k,j}$ denotes the response from the jth subject in the kth prior anchoring question. To control for the large skew in the estimates, we performed a log transformation of the raw data, making the support for each Θ_k stretch from negative infinity to infinity. We can assume a normal prior representation for each subject in each anchoring question so that

$$x_{k,j} \sim \text{Prior}(\theta_{k,j}, \delta_{k,j}) \\ \sim \mathcal{N}(\theta_{k,j}, \delta_{k,j}). \tag{B.1}$$

To further constrain the model we built the confidence judgments directly into the estimates for each of the $\delta_{k,j}$s, as discussed above. We first collapsed nearby confidence judgments into new bins, reducing the number of possible confidence responses from 13 to 6. For example, confidence responses of "1" and "2" were reassigned to the judgment "1" and "3" and "4" were reassigned to the judgment "2", and so on. The final bin, "6", contained judgments of "11", "12", and "13".

We denote the confidence response from the jth subject in the kth prior anchoring question as $z_{k,j}$ which ranges from $\{1, 2, \ldots, 6\}$. We group all subjects eliciting the same confidence judgment together by assuming that they have the same degree of prior uncertainty, although they may have different mean parameters $\theta_{k,j}$. Thus, for each confidence judgment $c \in \{1, 2, \ldots, 6\}$ in the kth question, we assign a separate prior uncertainty parameter $\beta_{k,c}$, so that

$$\delta_{k,j} = \beta_{k,z_{k,j}}. \tag{B.2}$$

To avoid any undue constraint on the estimates for $\beta_{k,c}$, we did not specify any hierarchical structure such as the ones described below; instead, we allowed each $\beta_{k,c}$ to vary freely.

Eqs. (B.1) and (B.2) specify that the data from the prior condition comes from the prior representation governed by the parameters $\theta_{k,j}$ and $\delta_{k,j} = \beta_{k,z_{k,j}}$.[11] Thus, the influence of the data $x_{k,j}$ on the parameters $\theta_{k,j}$ and $\delta_{k,j}$ can be described by the equation

$$f(\theta_{k,j}, \delta_{k,j} | x_{k,j}) = \text{Prior}(x_{k,j} | \theta_{k,j}, \delta_{k,j}). \tag{B.3}$$

B.2. The information function

As we explained above, AIM assumes that an individual evaluates how consistent the anchor is with the prior beliefs. We assume that upon the presentation of an anchor α_k in the kth question, individuals first determine a consistency value $\psi_{k,j}$ by integrating their prior representation up to the given anchor value, or

$$\psi_{k,j} = \int_{-\infty}^{\alpha_k} \text{Prior}(\Theta_k | \theta_{k,j}, \delta_{k,j}) d\Theta_k.$$

Once computed, this degree of consistency measure $\psi_{k,j}$ is then used to specify the anchor influence parameter $\lambda_{k,j}$ by evaluating

$$\lambda_{k,j} = B(\psi_{k,j} | P, w) \lambda^k_{\max}, \tag{B.4}$$

[11] Although several nodes in the model are deterministic, we will refer to them as parameters for simplicity. For example, the node $\delta_{k,j}$ is actually determined by the equation $\delta_{k,j} = \beta_{k,z_{k,j}}$, where $\beta_{k,z_{k,j}}$ is the unknown parameter.

where $w = \{1, \xi_j, \xi_j, 1\}$, $B(\psi_{k,j}|P, w)$ is evaluated by Eq. (3), and $\lambda^k_{\max}$ is the upper bound constraint on λ in the kth question. We let ξ vary across participants and assume that these parameters follow a common distribution such that

$$\xi_j \sim \mathcal{TN}(\xi^{(\mu)}, \xi^{(\sigma)}, 0, \infty), \tag{B.5}$$

where $\mathcal{TN}(a, b, c, d)$ represents a truncated normal distribution with mean a, standard deviation b, lower bound c and upper bound d.

The anchor influence parameter is then used to form the information function. Here, we again assume the information function is normal in form, so that

$$\text{Info}(\Theta_k|\alpha_k, \lambda_{k,j}) = \frac{1}{\sqrt{2\pi}\lambda_{k,j}} \exp\left[-\frac{(\Theta_k - \alpha_k)^2}{2\lambda^2_{k,j}}\right]. \tag{B.6}$$

B.3. The posterior representation

Once the anchor provides information about Θ_k, the participant updates their prior representation by the information function to arrive at the posterior representation

$$\text{Post}(\Theta_k|\theta_{k,j}, \delta_{k,j}, \lambda_{k,j}, \alpha_k) = \frac{\text{Prior}(\Theta_k|\theta_{k,j}, \delta_{k,j})\text{Info}(\Theta_k|\alpha_k, \lambda_{k,j})}{\int \text{Prior}(\Theta_k|\theta_{k,j}, \delta_{k,j})\text{Info}(\Theta_k|\alpha_k, \lambda_{k,j})d\Theta_k}.$$

For the posterior judgment $y_{k,j}$, we can write the influence of the data on the parameters $\theta_{k,j}, \delta_{k,j}$ and $\lambda_{k,j}$ as

$$f(\theta_{k,j}, \delta_{k,j}, \lambda_{k,j}|y_{k,j}, \alpha_k) = \text{Post}(y_{k,j}|\theta_{k,j}, \delta_{k,j}, \lambda_{k,j}, \alpha_k), \tag{B.7}$$

where α_k is assumed to be known and $\lambda_{k,j}$ is determined by Eq. (B.4).

Combining Eqs. (B.3) and (B.7), we can write the likelihood function as

$$L(\theta_{k,j}, \delta_{k,j}, \lambda_{k,j}|x_{k,j}, y_{k,j}, \alpha_k) = \text{Prior}(x_{k,j}|\theta_{k,j}, \delta_{k,j}) \times \text{Post}(y_{k,j}|\theta_{k,j}, \delta_{k,j}, \lambda_{k,j}, \alpha_k). \tag{B.8}$$

By assuming that the data $x_{k,j}$ and $y_{k,j}$ are independent and identically distributed from the data set x and y, respectively, we can express the likelihood function in Eq. (B.8) for the parameter set θ, δ, and λ as

$$L(\theta, \delta, \lambda|x, y, \alpha) = \prod_{k=1}^{K}\prod_{j=1}^{J}\text{Prior}(x_{k,j}|\theta_{k,j}, \delta_{k,j})\text{Post}(y_{k,j}|\theta_{k,j}, \delta_{k,j}, \lambda_{k,j}, \alpha_k), \tag{B.9}$$

where we assume that x and y are of the same dimensionality (namely, they are matrices of dimension $K \times J$). Because the free parameters in the model are actually θ, ξ, and β, it may be convenient to rewrite Eq. (B.9) in terms of these parameters, which is given by

$$L(\theta, \beta, \xi|x, y, z, \alpha) = \prod_{k=1}^{K}\prod_{j=1}^{J}\text{Prior}(x_{k,j}|\theta_{k,j}, \beta_{k,z_{k,j}}) \times \text{Post}(y_{k,j}|\theta_{k,j}, \beta_{k,z_{k,j}}, q_{k,j}, \alpha_k), \tag{B.10}$$

where

$$q_{k,j} = B\left(\int_{-\infty}^{\alpha_k} \text{Prior}(\Theta_k|\theta_{k,j}, \beta_{k,z_{k,j}})d\Theta_k|P, w = \{1, \xi_j, \xi_j, 1\}\right)\lambda^k_{\max}.$$

Note that this new expression makes the dependency on the observed confidence judgments z explicit.

Similar to the earlier version of our model used in Experiment 1 where we assumed a normal prior representation and a normal information function, we can combine Eqs. (B.1) and (B.6) to show that the posterior representation for these data also has a closed form expression. Specifically,

$$\text{Post}(\Theta_k|\theta_{k,j},\delta_{k,j},\lambda_{k,j},\alpha_k) = \frac{1}{\sqrt{2\pi}b_{k,j}} \exp\left[-\frac{(\Theta_k - a_{k,j})^2}{2b_{k,j}^2}\right],$$

which is a normal distribution with mean $a_{k,j}$ and standard deviation $b_{k,j}$, where

$$a_{k,j} = \frac{\frac{\theta_{k,j}}{\delta_{k,j}^2} + \frac{\alpha_k}{\lambda_{k,j}^2}}{\frac{1}{\delta_{k,j}^2} + \frac{1}{\lambda_{k,j}^2}}, \text{ and}$$

$$b_{k,j} = \left(\frac{1}{\delta_{k,j}^2} + \frac{1}{\lambda_{k,j}^2}\right)^{-1/2}. \tag{B.11}$$

Thus, the jth response from the kth anchoring question is assumed to arise from the distribution given by

$$\begin{aligned} y_{k,j} &\sim \text{Post}(\theta_{k,j},\delta_{k,j},\lambda_{k,j},\alpha_k) \\ &\sim \mathcal{N}(a_{k,j}, b_{k,j}) \end{aligned}$$

B.4. Estimating the model parameters

To fit the model to the data in our Experiment 2, we assumed that the prior representations and the information functions were normal in form. Thus, Eq. (B.9) becomes

$$L(\theta,\delta,\lambda|x,y,\alpha) = \prod_{k=1}^{K}\prod_{j=1}^{J}\phi(x_{k,j}|\theta_{k,j},\delta_{k,j})\phi(y_{k,j}|a_{k,j},b_{k,j}), \tag{B.12}$$

where $a_{k,j}$ and $b_{k,j}$ are determined by Eq. (B.11). With the likelihood function in hand, one can employ standard optimization routines to maximize Eq. (B.9) – or Eq. (B.12) when the appropriate assumptions are made – to obtain maximum likelihood estimates of the parameters θ, β, and ξ.

B.4.1. Bayesian inference

When fitting the within-participant version of AIM in the Bayesian framework, one must specify hierarchical structures in the form of prior distributions as well as specify prior distributions for these hierarchical parameters. In our model fitting, we did not specify any hierarchical structure for $\beta_{k,c}$, and instead allowed each $\beta_{k,c}$ to vary freely by specifying a uniform prior so that

$$\beta_{k,c} \sim CU[0, 20],$$

where $CU[a,b]$ denotes a continuous uniform distribution on the closed interval $[a,b]$. This specification made the prior density for β

$$\pi(\beta) = \prod_{k=1}^{K}\prod_{c=1}^{C}\frac{1}{20} = \left(\frac{1}{20}\right)^{KC}$$

By contrast, we constructed a hierarchical structure for the prior mean parameters by specifying that the set of prior means in the kth question, denoted $\theta_{k,1:J}$, come from a common normal distribution such that

$$\theta_{k,1:J} \sim \mathcal{N}\left(\theta_k^{(\mu)}, \theta_k^{(\sigma)}\right).$$

Thus, the prior density of $\theta_{k,1:J}$ is

$$\pi(\theta_{k,1:J}) = \prod_{j=1}^{J}\phi\left(\theta_{k,j}|\theta_k^{(\mu)}, \theta_k^{(\sigma)}\right),$$

and the prior density of θ is

$$\pi(\theta) = \prod_{k=1}^{K}\prod_{j=1}^{J}\phi\left(\theta_{k,j}|\theta_k^{(\mu)}, \theta_k^{(\sigma)}\right),$$

For the anchor sensitivity parameters ξ_j, we assumed a common truncated normal distribution, given by Eq. (B.5), such that the prior density of ξ is

$$\pi(\xi) = \prod_{j=1}^{J}\text{TN}\left(\xi_j|\xi^{(\mu)}, \xi^{(\sigma)}, 0, \infty\right),$$

where $\text{TN}(x|a,b,c,d)$ represents the truncated normal density with mean a, standard deviation b, lower bound c and upper bound d, evaluated at the location x.

We must now specify the prior distributions for the hyperparameters. For the means $\theta_k^{(\mu)}$ and standard deviations $\theta_k^{(\sigma)}$, we assumed mildly informative priors such that

$$\theta_k^{(\mu)} \sim \mathcal{N}(4, 10)$$
$$\theta_k^{(\sigma)} \sim \Gamma(1, 1),$$

for each $k \in \{1, \ldots, 6\}$. For the mean $\xi^{(\mu)}$ and hyper standard deviation $\xi^{(\sigma)}$, we assumed

$$\xi^{(\mu)} \sim \mathcal{TN}(10, 20, 0, \infty)$$
$$\xi^{(\sigma)} \sim \Gamma(1, 1).$$

Combining the likelihood function in Eq. (B.10) and each of the priors listed above, we can write the joint posterior distribution as

$$\pi(\eta|x, y, z, \alpha) \propto L(\theta, \xi, \beta|x, y, z, \alpha)\prod_{k=1}^{K}\left[\pi\left(\theta_k^{(\mu)}\right)\pi\left(\theta_k^{(\sigma)}\right)\right] \times \pi\left(\xi^{(\mu)}\right)\pi\left(\xi^{(\sigma)}\right)\pi(\theta)\pi(\xi)\pi(\beta), \tag{B.13}$$

where we use η to represent the set of parameters of interest, namely $\eta = \{\theta, \xi, \beta, \theta^{(\mu)}, \theta^{(\sigma)}, \xi^{(\mu)}, \xi^{(\sigma)}\}$. Eq. (B.13) is a general equation for fitting the model under the constraints we used to fit the model to the data from Experiment 2. There are numerous alternative modeling approaches, but the derivation of the joint posterior distribution would be similar to the one derived here.

To sample from the joint posterior in Eq. (B.13), we implemented a blocked version of the DE-MCMC algorithm (see Turner et al., 2013, for details). While we do not present the conditional distributions here, our blocks consisted of the pair $\left(\xi^{(\mu)}, \xi^{(\sigma)}\right)$, each kth pair $\left(\theta_k^{(\mu)}, \theta_k^{(\sigma)}\right)$ taken separately, each kth set of $\beta_{k,1:C}$, and each jth set $(\theta_{1:K,j}, \xi_j)$.

Table C.3
Key variable definitions for implementing the JAGS code for the between-participant version of AIM.

Variable	Description
n.post	Total number of observations in the "posterior" conditions
n.cond	Number of "posterior" conditions
n.prior	Number of observations in "prior" condition
mu	Value for the mean parameter on the prior for θ
alpha	Vector containing the values of the anchors used in each "posterior" condition
prior	A vector containing observations from the "prior" condition
post	Matrix containing (column 1) the observed "posterior" data, and (column 2) corresponding condition number

Appendix C. JAGS code for implementing the between-participant version of AIM

Below is JAGS code for implementing the between-participant version of AIM. In addition, a downloadable package is available on each of the authors' websites that contains (1) the JAGS code below, (2) R code for executing the JAGS code on Question 1, and (3) the data files from Experiment 1. Finally, Table C.3 contains a list of the variables that need to be passed to JAGS for implementation.

```
# Inferring the parameters of the between-participant version of AIM
model{
  # Prior Data
  for (i in 1:n.prior){
      prior[i] ~ dnorm(theta,1/delta^2)
  }
  # Posterior Data
  for (i in 1:n.post){
      den[i] <- ((1/delta^2) + (1/lambda[post[i,2]]^2))
      b[i] <- den[i]^(-1/2)
      a[i] <- ((theta/delta^2) + (alpha[post[i,2]]/lambda[post[i,2]]^2))
      / den[i]
      post[i,1] ~ dnorm(a[i],1/b[i]^2)
  }
  # Priors
  theta ~ dnorm(mu,1/10^2)
  delta ~ dgamma(1,1)
  for(k in 1:n.cond){
    lambda[k] ~ dgamma(1,1)
  }
}
```

Appendix D. JAGS code for implementing the within-participant version of AIM

Below is JAGS code for implementing the within-participant version of AIM. In addition, a downloadable package is available on each of the authors' websites that contains (1) the JAGS code below, (2) R code for executing the JAGS code on the data reported in Experiment 2, and (3) the data files from Experiment 2. Finally, Table D.4 contains a list of the variables that need to be passed to JAGS for implementation.

Table D.4
Key variable definitions for implementing the JAGS code for the within-participant version of AIM.

Variable	Description
n.questions	Total number of questions
n.subjects	Total number of subjects
n.confcat	Total number of confidence categories
lambda.max	Vector containing the maximum value of λ
alpha	Vector containing the values of the anchors used in each "posterior" condition
z	Matrix containing confidence judgments
prior	Matrix containing estimates in "prior" condition
post	Matrix containing estimates in "posterior" condition
P	Matrix containing the set of control points in Bezier curve

```
# Inferring the parameters of the within-participant version of AIM
model{
  for(k in 1:n.questions){
    for(j in 1:n.subjects){
      # delta[j,k] <- beta[k,z[j,k]]
      prior[j,k] ~ dnorm(theta[j,k], 1/beta[k,z[j,k]]^2)
      psi[j,k] <- pnorm(alpha[k],theta[j,k],1/beta[k,z[j,k]]^2)
      #n=3; Bezier curve calculated manually below.
      b0[j,k] <- (1-psi[j,k])^(3)
      b1[j,k] <- 3 * psi[j,k] * (1-psi[j,k])^2
      b2[j,k] <- 3 * psi[j,k]^2 * (1-psi[j,k])
      b3[j,k] <- psi[j,k]^(3)
      #w=[1, xi[j], xi[j], 1]
      den[j,k] <- b0[j,k]*1 + b1[j,k]*xi[j] + b2[j,k]*xi[j] + b3[j,k]*1
      num[j,k] <- b0[j,k]*1*P[1,2] + b1[j,k]*xi[j]*P[2,2]
                 +b2[j,k]*xi[j]*P[3,2] + b3[j,k]*1*P[4,2]
      B.psi[j,k] <- num[j,k]/den[j,k]
      lambda[j,k] <- B.psi[j,k]*lambda.max[k]
      den2[j,k] <- ((1/beta[k,z[j,k]]^2) + (1/lambda[j,k]^2))
      b[j,k] <- den2[j,k]^(-1/2)
      a[j,k] <- ((theta[j,k]/beta[k,z[j,k]]^2) + (alpha[k]/lambda[j,k]^2))
      / den2[j,k]
      post[j,k] ~ dnorm(a[j,k],1/b[j,k]^2)
    }}
  # priors
  for(k in 1:n.questions){
    for(j in 1:n.subjects){
      theta[j,k] ~ dnorm(theta.mu[k], 1/theta.sigma[k]^2)
    }}
  for(k in 1:n.questions){
  for(c in 1:n.confcat){
    beta[k,c] ~ dunif(0,20)
  }}
  for(j in 1:n.subjects){
    xi[j] ~ dnorm(xi.mu,1/xi.sigma^2) T(0,)
  }
  for(k in 1:n.questions){
    theta.mu[k] ~ dnorm(4,1/10^2)
    theta.sigma[k] ~ dgamma(1,1)
  }
  xi.mu ~ dnorm(10,1/20^2) T(0,)
  xi.sigma ~ dgamma(1,1)
}
```

References

Adaval, R., & Monroe, K. B. (2002). Automatic construction and use of contextual information for product and price evaluations. *Journal of Consumer Research, 28*, 572–588.

Allcott, H., & Mullainathan, S. (2010). Behavioral science and energy policy. *Science, 327*, 1204–1205.

Anderson, J. R. (1990). *Cognitive psychology and its implications*. New York, NY: Worth Publishers.

Ariely, D., Loewenstein, G., & Prelec, D. (2003). Coherent arbitrariness: Stable demand curves without stable preferences. *The Quarterly Journal of Economics, 118*, 73–105.

Baker, C. L., Tenenbaum, J. B., & Saxe, R. R. (2007). Goal inference as inverse planning. In *Proceedings of the 29th annual meeting of the cognitive science society*. .

Bhatia, S., & Chaudhry, S. J. (2013). The dynamics of anchoring in bidirectional associative memory networks. In *Proceedings of the 35th annual conference of the cognitive science society*. .
Birnbaum, M. H., & Zimmermann, J. M. (1998). Buying and selling prices of investments: Configural weight model of interactions predicts violations of joint independence. *Organizational Behavior and Human Decision Processes, 74*, 145–187.
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.
Brewer, N. T., & Chapman, G. B. (2002). The fragile basic anchoring effect. *Journal of Behavioral Decision Making, 7*, 223–242.
Brown, S., & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review, 112*, 117–128.
Brown, S., & Heathcote, A. (2008). The simplest complete model of choice reaction time: Linear ballistic accumulation. *Cognitive Psychology, 57*, 153–178.
Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*, 3–5.
Carlson, B. W. (1990). Anchoring and adjustment in judgments under risk. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 665–676.
Cervone, D., & Peake, P. K. (1986). Anchoring, efficacy, and action: The influence of judgmental heuristics on self-efficacy judgments and behavior. *Journal of Personality and Social Psychology, 50*, 492–501.
Chapman, G. B., & Bornstein, B. H. (1996). The more you ask for, the more you get: Anchoring in personal injury verdict. *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.
Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation and the construction of value. *Organizational Behavior and Human Decision Processes, 79*, 115–153.
Chapman, G. B., & Johnson, E. J. (2002). Incorporating the irrelevant: Anchors in judgment of belief and value. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 120–138). Cambridge, England: Cambridge University Press.
Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences, 10*, 335–344.
Critcher, C. R., & Gilovich, T. (2008). Incidental environmental anchors. *Journal of Behavioral Decision Making, 21*, 241–251.
Englich, B. (2008). When knowledge matters: Differential effects of available knowledge in standard and basic anchoring. *European Journal of Social Psychology, 38*, 896–904.
Englich, B., & Mussweiler, T. (2001). Sentencing under uncertainty: Anchoring effects in the courtroom. *Journal of Applied Social Psychology, 31*, 1535–1551.
Englich, B., Mussweiler, T., & Strack, F. (2006). Playing dice with criminal sentences: The influence of irrelevant anchors on expers' judicial decision making. *Personality and Social Psychology Bulletin, 32*, 188–200.
Epley, N. (2004). A tale of tuned decks? Anchoring as accessibility and anchoring as adjustment. In D. J. Koehler & N. Harvey (Eds.), *The Blackwell handbook of judgment and decision making* (pp. 240–256). Oxford, UK: Blackwell.
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.
Epley, N., & Gilovich, T. (2006). The anchoring and adjustment heuristic: Why adjustments are insufficient. *Psychological Science, 17*, 311–318.
Frederick, S., & Mochon, D. (2012). A scale distortion theory of anchoring. *Journal of Experimental Psychology: General, 141*, 124–133.
Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics, 40*, 35–42.
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.
Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis*. New York, NY: Chapman and Hall.
Gilbert, D. T. (2002). Inferential correction. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 167–184). Cambridge, England: Cambridge University Press.
Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge, England: Cambridge University Press.
Goldstein, W. M., & Einhorn, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological Review, 94*, 236–254.
Green, D., Jacowitz, K. E., Kahneman, D., & McFadden, D. (1998). Referendum contingent valuation, anchoring, and willingness to pay for public goods. *Resource and Energy Economics, 20*, 85–116.
Grice, H. R. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58). New York: Academic Press.
Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition. In R. Sun (Ed.), *Cambridge handbook of computational cognitive modeling*. Cambridge University Press.
Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological Science, 17*, 725–743.
Griffiths, T. L., & Tenenbaum, J. B. (2011). Predicting the future as Bayesian inference: People combine prior knowledge with observations when estimating duration and extent. *Journal of Experimental Psychology: General, 140*, 725–743.
Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology, 24*, 1–55.
Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin, 21*, 1161–1167.
Johnson, J. G., & Busemeyer, J. R. (2005). A dynamic, stochastic, computational model of preference reversal phenomena. *Psychological Review, 112*, 841–861.
Johnson, E. J., & Goldstein, D. G. (2003). Do defaults save lives? *Science, 302*, 1338–1339.
Johnson, E. J., & Schkade, D. A. (1989). Bias in utility assessments: Further evidence and explanations. *Management Science, 35*, 406–424.

Kahneman, D., & Tversky, A. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.
Klayman, J., & Ha, Y. W. (1991). Confirmation, disconfirmation, and information in hypotheses testing. *Psychological Review, 94*, 211–228.
Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*, 22–44.
Lee, M. D. (2004). A Bayesian analysis of retention functions. *Journal of Mathematical Psychology, 48*, 310–321.
Lee, M. D. (2008). Three case studies in the Bayesian analysis of cognitive models. *Psychonomic Bulletin and Review, 15*, 1–15.
Lee, M. D., & Wagenmakers, E.-J. (2010). A Course in Bayesian graphical modeling for cognitive science. Available from <http://users.fmg.uva.nl/ewagenmakers/BayesCourse/BayesBook.pdf>; last downloaded February 26, 2010.
Lewandowsky, S., Griffiths, T. L., & Kalish, M. L. (2009). The wisdom of individual: Exploring peoples knowledge about everyday events using iterated learning. *Cognitive Science, 33*, 969–998.
Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology, 89*, 46–55.
Lieder, F., Griffiths, T. L., & Goodman, N. D. (2012). Burn-in, bias, and the rationality of anchoring. In *Advances in Neural Information Processing Systems* (pp. 2699–2707).
Liu, C. C., & Aitkin, M. (2008). Bayes factors: Prior sensitivity and model generalizability. *Journal of Mathematical Psychology, 52*, 362–375.
McElroy, T., & Dowd, K. (2007). Susceptibility to anchoring effects: How openness-to-experience influences responses to anchoring cues. *Judgment and Decision Making, 2*, 48–53.
Mochon, D., & Frederick, S. (2013). Anchoring in sequential judgments. *Organizational Behavior and Human Decision Processes, 122*, 69–79.
Mussweiler, T. (2001). The durability of anchoring effects. *European Journal of Social Psychology, 31*, 431–442.
Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review, 110*, 472–489.
Mussweiler, T., & Englich, B. (2005). Subliminal anchoring: Judgmental consequences and underlying mechanisms. *Organizational Behavior and Human Decision Processes, 98*, 133–143.
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.
Mussweiler, T., & Strack, F. (2000). The use of category and exemplar knowledge in the solution of anchoring tasks. *Journal of Personality and Social Psychology, 78*, 1038–1052.
Mussweiler, T., & Strack, F. (2001). Consider the impossible: Explaining the effects of implausible anchors. *Social Cognition, 19*, 145–160.
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.
Navarro-Martinez, D., Salisbury, L. C., Lemon, K. N., Stewart, N., Matthews, W. J., & Harris, A. J. L. (2011). Minimum required payment and supplemental information disclosure effects on consumer debt repayment decisions. *Journal of Marketing Research, 48*, S60–S77.
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*, 44–64.
Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*, 39–57.
Oppenheimer, D. M., LeBoeuf, R. A., & Brewer, N. T. (2008). Anchors aweigh: A demonstration of cross-modality anchoring and magnitude priming. *Cognition, 106*, 13–26.
Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5*, 411–419.
Petty, R. E., & Cacioppo, J. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.). *Advances in experimental social psychology* (Vol. 19, pp. 123–205). New York: Academic Press.
Plous, S. (1989). Thinking the unthinkable: The effect of anchoring on likelihood estimates of nuclear war. *Journal of Applied Social Psychology, 19*, 67–91.
Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing*. .
Quattrone, G. A. (1982). Overattribution and unit formation: When behavior engulfs the person. *Journal of Personality and Social Psychology, 42*, 593–607.
Quattrone, G. A., Lawrence, C. P., Finkel, S. E., & Andrus, D. C. (1984). Explorations in anchoring: The effects of prior range, anchor extremity, and suggestive hints (unpublished manuscript).
Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59–108.
Robert, C. P., & Casella, G. (2004). *Monte Carlo statistical methods*. New York, NY: Springer.
Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review, 4*, 734–760.
Schwarz, N. (1994). Judgments in a social context: Biases, shortcomings, and the logic of conversation. In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 125–162). San Diego: Academic Press.
Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science, 237*, 1317–1323.
Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. London: Chapman & Hall.
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*, 917–932.
Simonson, I., & Drolet, A. (2004). Anchoring effects on consumers' willingness-to-pay and willingness-to-accept. *Journal of Consumer Research, 31*, 681–690.
Slovic, P. (1967). The relative influence of probabilities and payoffs upon perceived risk of a gamble. *Psychonomic Science, 9*, 223–224.
Slovic, P., & Lichtenstein, S. (1968). Relative importance of probabilities and payoffs in risk taking. *Journal of Experimental Psychology Monograph, 78*, 1–18.

Smith, A. R., & Windschit, P. D. (2015). Resisting anchoring effects: The roles of metric and mapping knowledge. *Memory & Cognition, 43*, 1071–1084.

Smith, A. R., Windschitl, P. D., & Bruchmann, K. (2013). Knowledge matters: Anchoring effects are moderated by knowledge level. *European Journal of Social Psychology, 43*, 97–108.

Stewart, N. (2009). The cost of anchoring on credit card minimum payments. *Psychological Science, 20*, 39–41.

Steyvers, M., Griffiths, T. L., & Dennis, S. (2006). Probabilistic inference in human semantic memory. *Trends in Cognitive Sciences, 10*, 327–334.

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

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences, 10*, 309–318.

ter Braak, C. J. F. (2006). A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: Easy Bayesian computing for real parameter spaces. *Statistics and Computing, 16*, 239–249.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Treisman, M., & Williams, T. (1984). A theory of criterion setting with an application to sequential dependencies. *Psychological Review, 91*, 68–111.

Turner, B. M., Sederberg, P. B., Brown, S., & Steyvers, M. (2013). A method for efficiently sampling from distributions with correlated dimensions. *Psychological Methods, 18*, 368–384.

Turner, B. M., Van Zandt, T., & Brown, S. D. (2011). A dynamic, stimulus-driven model of signal detection. *Psychological Review, 118*, 583–613.

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

Urminsky, O., & Goswami, I. (2015). In search of optimally effective defaults. Working paper.

Usher, M., & McClelland, J. L. (2001). On the time course of perceptual choice: The leaky competing accumulator model. *Psychological Review, 108*, 550–592.

van Rooijen, M. R., & Daamen, D. D. L. (2006). Subliminal anchoring: The effects of subliminally presented numbers on probability estimates. *Journal of Experimental Social Psychology, 42*, 380–387.

Vul, E., Goodman, N., Griffiths, T. L., & Tenenbaum, J. B. (2014). One and done? Optimal decisions from very few samples. *Cognitive Science, 38*(4), 599–637.

Wansink, B., Kent, R. J., & Hoch, S. J. (1998). An anchoring and adjustment model of purchase quantity decisions. *Journal of Marketing Research, 35*, 71–81.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12*, 129–140.

Wegener, D. T., Petty, R. E., Blankenship, K. L., & Detweiler-Bedell, B. (2010). Elaboration and numerical anchoring: Breadth, depth, and the role of (non-)thoughtful processes in anchoring theories. *Journal of Consumer Psychology, 20*, 5–16.

Wegener, D. T., Petty, R. E., Detweiler-Bedell, B., & 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.

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*, 74–85.

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.

Wixted, J. T. (1990). Analyzing the empirical course of forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 927–935.

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.

Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences, 10*, 301–308.

Zhang, Y., Li, Y., & Zhu, T. (2011). How multiple anchors affect judgment: Evidence from the Lab and eBay. Working paper.

Zhang, Y. C., & Schwarz, N. (2013). The power of precise numbers: A conversational logic analysis. *Journal of Experimental Social Psychology, 49*, 944–946.