Fall 12-17-2014

A Sticky Space Model for Explanation and Individuation of Anchoring Effects

Robert Hatcher

A STICKY SPACE MODEL FOR EXPLANATION AND INDIVIDUATION OF ANCHORING EFFECTS

by

ROBERT HATCHER

Under the Direction of Neil Van Leeuwen

ABSTRACT

Current explanations for anchoring phenomena seem to be unable to account for the diversity of effects found by 40 years of research. Additionally, the theories do not have much to say about the processes that make anchors so resilient to modification. I argue that by focusing on the mechanisms involved in spatial representation, we can account for most anchoring effects which have spatial components.

INDEX WORDS: Spatial representation, Anchoring, Cognitive bias, Rationality

A STICKY SPACE MODEL FOR EXPLANATION AND INDIVIDUATION OF ANCHORING EFFECTS

by

ROBERT HATCHER

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2014

A STICKY SPACE MODEL FOR EXPLANATION AND INDIVIDUATION OF ANCHORING EFFECTS

by

ROBERT HATCHER

Committee Chair: Neil Van Leeuwen

Committee: Eddy Nahmias

Daniel Weiskopf

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

December 2014

DEDICATION

To Gail Hatcher

ACKNOWLEDGEMENTS

First and foremost I would like to thank Neil Van Leeuwen, Eddy Nahmias, and Dan Weiskopf for the helpful comments in developing this thesis. I would also like to thank Barbra Sonnies and Gail Hatcher for their editorial prowess. Finally, I would like to thank Elisabeth Proctor, Rachel Kaye, Harvey Kupferberg, and the rest of my friends, family, and colleagues who have been so generous for their support during the process.

TABLE OF CONTENTS

1 INTRODUCTION

Forty years of research into the phenomenon of anchoring suggests that information we generate (as opposed to recall from memory), is highly susceptible to being biased towards information recently under consideration. We are biased even when the previous information is considered irrelevant to our task. Anchoring is a robust phenomenon found in many different areas of our life. Anchoring effects have been found in everything from trivia knowledge, performance evaluations, price evaluations, and even how well we can draw lines on a chalkboard. The range of domains anchoring effects can be found in should make us wonder if all anchoring phenomenon are the result of a singular process. Most researchers who study anchoring specifically claim that anchoring effects might stem from multiple mental processes. For example, anchoring effects have been hypothesized to stem from numerical or conceptual priming (Jacowitz and Kahneman, 1995; Wong and Kwong, 2000; Oppenheimer, LeBoeuf, and Brewer, 2008), semantic priming (Chapman and Johnson, 1999; Mussweiler and Strack, 2001; Strack and Mussweiler, 1997), shifts in response scales (S. W. Frederick and Mochon, 2011), and effortful adjustment (Epley and Gilovich, 2005; Wegener et al., 2010). If anchoring effects come from many different mental processes, then we want to know what processes govern which types of anchoring effects. I call this the *individuation problem*. The *individuation problem* says that we need a principled way to determine which anchoring phenomena are correctly grouped together. In other words, we need a principled way of determining what kind of anchoring process will be involved in any given anchoring example. Individuating anchoring phenomenon means having a clear idea of the mechanisms that can cause each type, so we want theories which provide such mechanisms.

I will argue that some anchoring phenomena are a consequence of what I call a "sticky" representational space. What it means for our representational space to be "sticky" will fleshed out in more detail later. For the moment, we can think of a sticky space as being one that resists changes in the scale of the space which we represent objects spatially. It might seem strange, at first, to think that

anchoring should be found in spatial tasks given how well we perform tasks like navigation, picking up objects, and judging whether a shirt has shrunk in the drier. If, for example, thinking about your monthly budget affected subsequent judgments we might think that the number of car accidents should be much higher than they already are.

There is a difference between anchoring, as it is typically brought about, and actions like driving on the highway. When we are driving we have reliable (though not infallible) visual information which we can refer to when we need to determine if it is safe to change lanes. Anchoring tasks usually involve questions whose answers we are going to be very uncertain about. For example, a common task in anchoring studies is to estimate the length of the Mississippi River. While many of us have some idea of the river's length (e.g. the river extends almost the full longitudinal height of the U.S.), we don't have an exact figure in mind. Anchoring studies have shown that when we first consider whether the river is longer (or shorter) than, say, 70 miles, our absolute estimates of length wind up being closer to 70 miles than otherwise might have occurred. If people's estimates tend to be closer to that 70 mile value, it is claimed, they have incorporated a value they know to be false into their estimation. The question is why do we incorporate the 70 mile value?

I will argue that when we perform estimation tasks that involve spatial dimensions, we generate our responses by the same basic process that we generate estimates of distance and speed of objects on the highway. Roughly stated, my argument is that we use the adaptive scaling of our representational system to generate information about the world. During navigation and object tracking in the 'real world', the scale of that grid is very stable because it is set by certain types of objects in our environment. Having a stable, but adaptable, scale to the spatial grid is advantageous because it allows for a range of behavior that would otherwise not be possible. However, when we engage in imagined or hypothetical situations the spatial grid becomes disconnected from the objects that usually set its size. Therefore, the scale of the spatial grid is subject to alterations and distortions. When we try to use the

spatial grid to estimate spatial values in communicable terms, those distortions transfer to our response values. Those (distorted) values account, I claim, for anchoring effects. Since not all anchoring effects need to use the spatial grid of our representational system, the model I present is limited. That limitation, however, is actually a benefit because it allows us to differentiate *spatial anchoring* from anchoring phenomenon that employ numerical, temporal, or linguistic mechanisms.

The structure of this paper can be broken into two main parts. The first three sections will focus on anchoring effects. Section 2 will outline a few examples of anchoring with spatial components. This section is designed to provide the empirical findings which we will later use to assess theories. In section 3 I will put forth a list of desiderata which we want a full explanation of anchoring to incorporate. These desiderata will also provide us with a framework which allows us to solve the *individuation* problem. Then, in section 4, I will examine the major theories of anchoring to see how well they capture our desiderata and, thereby, properly individuate anchoring phenomenon.

The second part of the paper will be used to argue for what I will call the Sticky Space model of representation. In section 5, I begin by supporting three claims. The first claim is that we have evidence of a spatial grid which has a variable, but 'sticky', scale which we use to act in the world. I will define 'sticky' in more depth later, but it roughly means that the scale of our spatial grid resists change. Second, the scale is generally stable and information we glean from it is reliable. Third, there is a great benefit to having a scale which is sticky because it allows us to track the location of objects in the world as well as our relationship to those objects.

Section 6 will be devoted to showing how imaginative and hypothetical scenarios differ from 'real world' scenarios with regard to the spatial grid. I will argue that when we actively consider anchor values, the grid becomes less stable because it lacks the space defining objects which normally set the scale. When we begin representing objects on the spatial grid, the scale is subject to alterations and distortions as we try to balance competing elements of the imagined scene. Those distortions in the

scale distort the information which we generate from that scale (length, width, etc). I then map the

Sticky Space model onto several anchoring studies to show how the model accounts for anchoring

effects.

In section 7 I address the possible impact the Sticky Space model has on the 'Rationality

Wars'. I will argue that the underspecification of the major theories has led us to mischaracterize

anchoring effects. While I do not doubt that some anchoring effects are the result of 'irrational'

processes, I argue that this is not the case for *spatial anchoring*. Finally, I will conclude the paper by

looking at how the theory might be extended to other domains as well as areas of further research.

2   Empirical Results of Anchoring Research

For over 40 years, researchers have been investigating a phenomenon known as anchoring. Traditionally, anchoring effects have been described as the incorporation of prior (often irrelevant) information into subsequent judgments. When anchoring occurs, our subsequent judgments are biased toward the anchor value. It should be noted that comparing study participant's responses to the factual answer is not particularly important in assessing anchoring effects. The questions used in anchoring studies are often those which people are not likely to know the answer. Therefore, factually inaccurate answers should be expected. What *is* interesting about anchoring is that our responses differ when we first consider an anchor value. Anchoring effects have been found in estimates of various factual information like sizes of objects, distances between cities, the year certain events happened, and demographic information (Jacowitz and Kahneman, 1995). Anchoring effects are also found price estimates (Northcraft and Neale, 1987), value comparisons (Chapman and Johnson, 1999), performance evaluations (Switzer and Sniezek, 1991), negotiations (Galinsky, Mussweiler, and Medvec, 2002), causal attributions (Quattrone, 1982), and judicial sentencing (Englich and Mussweiler, 2001). Since the project I am engaging in here is concerned with anchoring effects with spatial components, I will limit the examples I use to those with estimates in length, size, etc. I will go into more depth about why such a limitation is appropriate in sections 3 and 4.

Before proceeding it will be useful to make a clarification about anchoring. There is a distinction in the literature between basic anchoring and traditional anchoring. Traditional anchoring is a two step paradigm. In the first step subjects are asked to evaluate whether an entity (the anchor) is smaller or larger than a particular value (the anchor value). In the second step, subjects are asked to estimate a point value for the target. Basic anchoring dispenses with the first step of the traditional paradigm. In basic anchoring, information about the anchor is made available to the subject but the subject is not specifically asked to evaluate anything. For example, subliminal anchor values have been

found to influence our subsequent judgments (Mussweiler and Englich, 2005; Wilson et al., 1996). There is some dispute about whether basic anchoring occurs (Brewer and Chapman, 2002), but when it is found the effects tend to be smaller and are more fragile (Critcher and Gilovich, 2008; Wilson et al., 1996). The important difference between basic anchoring and traditional anchoring is that subjects explicitly consider the anchor value in traditional anchoring. Explicitly considering the anchor value might cause the processes underlying anchoring effects to differ from those when the value is not explicitly considered. Since the theories I look at in section 3 all are based upon the traditional anchoring paradigm, I will not provide examples of basic anchoring.

Jacowitz and Kahneman (1995) asked participants a range of questions about general factual information such as "What is the length of the Mississippi River?" and "What is the population of Chicago?" Since we are concerned with spatial elements we can focus on the first question. Study participants were broken into three groups: calibration, low anchor, and high anchor. The calibration group was asked to estimate the length of the Mississippi River without being provided an anchor value. The 15th and 85th percentile answers were then used as the anchor values in the small and large anchor groups. Participants who were first asked if the Mississippi River is longer or shorter than 70 miles (small anchor group) gave a median estimate of 300 miles. Those who were provided a large anchor (2,000 miles) gave a median estimate of 1,500 miles. Considering that the median estimate of the Mississippi River was 800 miles in the calibration group, the effectiveness of anchors can be significant.

Mussweiler and Strack (2000, study 4) aimed to test the limits of that effectiveness. They asked participants if the Mississippi River was longer or shorter than 30,000 miles (large anchor). The median response during the estimation task was 3,768 miles. So, while the result showed the general trend of larger anchors producing larger absolute estimates, it also shows that the effectiveness increases at a decreasing rate. One way this might be accomplished is by using categorical information about the target to generate boundary conditions for responses. Mussweiler and Strack's idea is that we use background

information about how long any river could be to limit how large (or perhaps small) a particular river should be.

The idea that we use background knowledge, to help us arrive at our estimations, gets further support by studies which show anchoring effects even when we do not have numerical comparisons to make. Frederick and Mochon (2011) asked participants to examine a list of animals ordered by real world weight and pick the animal whose weight was closest to 1,000 pounds. Half of the participants were first asked to estimate the weight of a wolf prior to this task. Researchers found that participants who estimated the weight of the wolf chose animals that were larger than participants who did not. The average real world weight of the animals chosen by those who did not do the anchoring task was 1,385lbs. The average real world weight of the animals chosen by the anchor group was 2,170lbs. In this task, we have to use background knowledge about how large the animals are in comparison to the wolf and then convert that information into a value in pounds. One question we might ask ourselves is whether the anchoring effect comes from alteration of the spatial information we have about other animals (in light of the wolf) or if the anchoring effect comes from having to convert that spatial information into the pounds scale. The authors advocate for the latter explanation, but there may be reason to think that the spatial information changes during anchoring.

LeBoeuf and Shafir (2006) hypothesized that if subjects tend to insufficiently adjust because of representational and perceptual constructs, then anchoring effects should be seen in the way people adjust physical objects. They asked participants to either extend or erase a line so that the resulting line would be 3.5 inches long. What they found is that when people had to erase a long line (the large anchor) they consistently made lines that were too long. When participants had to extend a smaller line (the small anchor) the lines were drawn too short. The effect occurs even when participants were allowed to examine a line of the correct length. Allowing participants to examine the line ensures that the effect was not driven by people not knowing what long 3.5 inches looks like (LeBoeuf and Shafir,

2006, 397). The result of the study seems to suggest that the spatial information recalled during the second stage of the study is altered in the presence of the anchor. The study, therefore, may have an impact on Mussweiler and Strack's findings that categorical information is used to limit the extent of anchor effectiveness. We might use categorical information to place boundaries on absolute estimates, but this study suggests that the categorical information may be subject to anchoring effects as well.

Oppenheimer et al (2008) successfully reproduced LeBoeuf and Shafir's findings in the line extension/erasure studies but also found that drawing lines affected estimations of length in other contexts. In one study, participants were first asked to replicate a short or long squiggly line. Participants drew their line directly underneath the referent, so there was no guessing about length or shape. Next, participants were asked to estimate the length of the Mississippi River. Participants who drew the short line gave significantly shorter estimations than participants who drew long lines. While the data was not conclusive, there was some evidence that the anchoring could be bi-directional. This means that numerical values might act as anchors when attempting to draw lines of a particular length.

Interestingly, Oppenheimer et al present evidence that shows anchoring effects might even cross modalities. In the previous study, the line drawing task and length estimation are along a single spatial dimension. In another study, Oppenheimer et al asked participants to draw a short/long line in the first phase, and then estimate the mean temperature of Honolulu in the second phase. Participants who drew short lines gave lower estimates of temperature than people who drew long lines. This evidence shows that spatial information, can act as an anchor in entirely different domains. That surprised Oppenheimer et al because "previous studies found reduced anchoring effects when targets and anchors were from incompatible domains or dimensions." (Oppenheimer, LeBoeuf, and Brewer 2008, 18)

The studies listed here are but a small sample of the research available on anchoring. I have tried to pull together studies which show some of the interesting features about anchoring effects

in spatial contexts. A number of the questions posed to subjects in anchoring studies get repeatedly

reused over time (e.g. the length of the Mississippi River or sizes of animals). The general results do not

differ from study to study, so we should conclude that the effects are robust. Anchor values are robust,

but do have limitations on the magnitude of their effect. Those limitations seem to be a result of spatial

information about categories of objects. It is an open question, at this point, just how that spatial

information is utilized and, more generally, why anchoring occurs. I will look at some of the major

theories of anchoring in section 4.

3   Desiderata and Explanation

Before turning to a critical analysis of various theories of anchoring, it will be helpful to have a list of features which we want an explanation of anchoring to have:

1) *The theory needs to be specific enough to derive the various phenomena.*

2) *The theory tells us why anchoring effects are persistent*

3) *The theory tells us why our response values are what they are*

4) *The theory unifies as many anchoring examples as possible.*

Desideratum (1) says that any explanation of anchoring needs to be more than a mere accounting of the various conditions which anchoring and the attendant phenomenon occur. What we ideally want is a theory which allows us to derive the phenomenon from a set of initial conditions. The theory should be able to provide a reason for why some conditions may elicit anchoring effects but other, generally similar, conditions fail to do so. For example, if mere exposure to numbers in a year format causes anchoring effects to appear but exposure to numbers in a zip-code format does not, we ideally want a theory which differentiates year/date mechanisms from zip-code mechanisms.

Desideratum (2) is designed to ensure that any theory of anchoring can account for why anchors are persistent. By persistent I mean two things:

(1) Anchor information (be it a numerical value or physical dimension) is persistent in that it is reused in subsequent judgments even though it does not seem relevant. We can call this *temporal persistence.*

(2) Anchoring is persistent in that it is hard to de-bias. We can call this *irreversibility persistence*.

To accommodate the first sense of the word persistent, a theory should be able to answer the question "Why do we not evaluate the target in isolation?" One reason may be that anchor information is considered relevant to the target question. For example, using the year America declared independence (the anchor value) to arrive at an estimate of when George Washington became president (the target) seems relevant because independence was a pre-condition of electing a president (Epley and Gilovich, 2001). Using a random number, however, does not seem like a reasonable approach to the task because that number is not useful.

To accommodate the second sense of the word persistent, a theory should be able to answer the question "Why is anchoring so hard to de-bias?" Telling subjects about anchoring effects does help mitigate the severity of anchoring, but the effects are still present. Ideally, a theory of anchoring will be able to explain why the anchor information contaminates our target evaluations even when we (presumably) try to ignore the anchor. Answering the question becomes more important when viewed in conjunction with temporal persistence. When subjects know about anchoring and they know the anchor information is not relevant to the target evaluation, why do they not say "Hey, I know I am susceptible to anchoring and the information I have just been presented with is not relevant to this new judgment. I had better completely ignore it!"

Desideratum (3) says that an explanation of anchoring phenomena should include some explanation for why we settle on the response values that we do. I do not mean that an explanation must be able to supply a mathematical model that allows us to compute participant's target values. Rather, we want a theory to tell us *why* the anchor value feels wrong and why our response value feels right. Consider that in each of the previous anchoring studies, we reject the anchor value as being too high or too low for the target. We can infer from this two things. First, something tells us that the anchor value is incorrect. If we did not have this 'intuition' then we would not make any adjustments in the first

place. Second, we have some sense that the target value is either higher or lower than the anchor. If we

did not have an idea of whether the anchor was too high or too low, then we would not know which way

to adjust from the anchor value. If we did not know which way to adjust, then we should expect

response values to be equally distributed above and below the anchor value. Finally, it seems like the

target value has to land in some range that we are willing to accept as 'probably close' or 'reasonable'.

That should tell us that we already have some value, or range of values, which we have assigned the

target. Ideally, a theory of anchoring will point to what processes generates that range of values and

incorporates the process into the theory.

Finally, Desideratum (4) merely says that we want a theory of anchoring from which we

can derive as many different examples of anchoring as possible. While it is common to talk about

anchoring as a single phenomenon, we should not automatically assume that all anchoring effects stem

from a single process. It would, of course, be ideal if a single theory could explain all anchoring

phenomena, but if different instances of anchoring stem from different processes, we should only expect

a theory to encompass anchoring effects at the process level of generality.

4   Anchoring Theories

Now that we have a set of anchoring examples and a set of desiderata which an explanation of anchoring needs to fulfill we can turn to an examination of the major anchoring theories. I have two main goals in this section. First, I want to evaluate how well three major theories of anchoring satisfy the desiderata from the last section. When evaluating the strengths and weaknesses of each theory I will focus my attention primarily on desiderata (1) - (3). I will wait until the end of the section to examine desideratum (4). I do this because without having a sense of how each theory deals with (1)-(3), we will not be able to fairly decide which (if any) experimental results the theory need to accommodate. We should not hold a theory accountable for explaining any example that is outside the range of phenomena the theory was meant to address. The second goal I have for this section is to raise the individuation problem. If the examples from section 1 are caused by different mental processes, the how can we decide which theory needs to explain which example? In other words, why should we think that any of the three theories I look at should be able to claim they carve anchoring at its joints? How the theories satisfy the first three desiderata will tell us how well they differentiate anchoring phenomena.

Anchor and Adjustment Theory

The oldest theory of anchoring phenomenon is the anchor-and-adjust model of Kahneman and Tversky. Anchor and adjustment theories all say that the anchor value is considered as a possible target value when evaluating the target. The anchor value is rejected and we adjust (up or down) from that value until we reach a value that we find reasonable for the target. In the classic example of anchor-and-adjustment comes from Tversky and Kahneman (1974). Subjects were asked to spin a roulette wheel with numerical values printed on it. Subjects were later asked if the number of African countries in the U.N. was higher or lower than the number shown on the wheel. According to Tversky and Kahneman, subjects use the number from the wheel (anchor value) as a place to start the

estimating process. In effect, what subjects do is say "I don't know the answer to the question, so this [anchor] value is as good a place to start as any." Subjects then adjust from that value until a reasonable value is found. Unfortunately, that adjustment is insufficient and response values wind up closer to the anchor value than they otherwise would.

Presumably, subjects in the Tversky and Kahneman study know that the number from wheel is uninformative in regards to the target evaluation. In cases where the subject knows the anchor is uninformative, it is not clear how an anchor-and-adjust model is able to account for why the anchor is persistent when they are told about anchoring effects (desideratum (2)). It would seem that informing subjects of anchoring effects should make them ignore the anchor value completely and, thereby, negate the effect. Additionally, it is not clear how the theory can accommodate desideratum (3). If subjects have no information about the target, why do the reject the anchor value and how would they decide which direction from the anchor value to adjust?

A key assumption of anchor-and-adjust theories is that subjects use background information to answer the first question in the standard anchoring paradigm. For example, during the first phase, subjects may be asked if the Mississippi River is longer or shorter than 70 miles. The thought was that subjects compare the anchor value to some concrete internal value, or range of values, to answer that question. The vector of the necessary adjustment during the second phase (when subjects are asked to estimate the length of the river) is provided by the internal value and we insufficiently adjust from the anchor value when specifying the point value to the target. The strategy of consulting background information seems plausible and even reasonable. For example, Epley and Gilovich (2001) asked subjects when George Washington was elected president of the United States. When prompted to verbally walk through the problem subjects would say "The United States declared its independence in 1776 and it probably took a few years to elect a president, so Washington was elected in...1779."(Epley and Gilovich, 2001, 392) Notice, however, that Epley and Gilovich did not provide an anchor value in the

question. That difference becomes important because research has shown that when provided an anchor, the internal value used for comparison during the first phase is subject to anchoring effects. In other words, before we even think about providing a point estimate of the target (e.g. the Mississippi River), we have already been biased towards the anchor (Jacowitz and Kahneman, 1995; Green et al., 1998). Recall that Jacowitz and Kahneman used the 15th and 85th percentile answers of the calibration group as the anchor values for the other groups. What Jacowitz and Kahneman found was that in the anchoring conditions, significantly more subjects claimed the anchor was higher (or lower) than the provided anchor value. For example, when asked whether the Mississippi River was longer or shorter than 2,000 miles (the large anchor), 37% of people responded that it was longer than 2,000 miles. By comparison, only 15% of the calibration group said that the Mississippi River was longer than 2,000 miles. If we had stable internal values for the Mississippi River, we should only expect 15% to say that given responses to non-anchored questions.

Even though the adjustment paradigm is no longer the dominant model, I think it is useful to pick out lessons we can learn from the failure of the anchor-and-adjust model. First, using background information (e.g. the year America declared independence or some internal representation of the Mississippi River) to guide our evaluation of the target seems plausible. In considering the answer to trivia questions, my phenomenological experience mirrors that of subjects in Epley and Gilovich. Second, that background information may be subject to anchoring. For example, in estimating the year George Washington became president, determining the time it took for the 'dust to settle' enough for an election to take place might be affected by anchors. The important point to pull from Jacowitz and Kahneman is that anchoring can take place before we attempt to provide an estimate of the target value.[1]

---

[1] Until Jacowitz and Kahneman, anchoring was thought to occur only during the estimation task. Because Jacowitz and Kahneman's results seemed to give evidence to the contrary, anchor-and-adjust theories were

Selective Accessibility Model

The dominant theory of anchoring, the Selective Accessibility Model (SAM), is an activation account championed by Chapman and Johnson (1999) and Mussweiler and Strack (2000). Activation accounts all roughly say that anchors selectively activate target information which is consistent with the anchor (see Chapman and Johnson, 1999; Mussweiler and Strack, 2007). More specifically, SAM states that when we when we evaluate a target in the presence of an anchor, we perform an automatic (Jacowitz and Kahneman, 1995; Chapman and Johnson, 1999) and biased search for semantic similarities to the anchor. The mental representation of the target is biased towards the anchor because of the increased availability of anchor specific features (Mussweiler and Strack, 2000). The search is considered biased because it focuses on similarities rather than relevant differences. We use information about how similar the target and anchor are in order for us to determine how close the target value should be to the anchor value (Chapman and Johnson, 1999, 121).

Evidence for the search for semantic similarity comes from studies which show subjects are much faster at recognizing words associated with the anchor (Chapman and Johnson, 1999; Mussweiler and Strack, 2000; Mussweiler and Strack, 2001; Mussweiler and Englich, 2005). For example, in a study done by Mussweiler and Strack (2000), participants were first asked to judge whether the mean temperature of Germany was higher or lower than an anchor value. Next, subjects were given a sting of words and asked to categorize them as either summer words or winter words. Participants who were provided the low anchor value were much faster at categorizing winter words (e.g. "cold" or "ski") than participants given high anchors. Conversely, those given high anchors were faster at identifying summer related words. Additionally, Chapman and Johnson (1999) show that prompting participants to consider different features of the anchor, anchoring effects can be mitigated.

---

deemed flawed. This study led to anchor-and-adjust theories falling out of favor. See Chapman and Johnson (1999) for further details.

The emphasis on semantic similarity causes several problems that relate to our desiderata. First, it is hard to see how we can derive the phenomenon when we locate the source of anchoring effects as semantic similarity. The problem arises because there are many similarities between objects which might be differ at various levels of representation. For example, when I ask you to estimate the size of a starling given a crow as the anchor, what counts as a similarity between the two birds at one level of abstraction might be completely different at another level. In the broadest sense both birds fly, both have calls, make nests, and (if seen from far away) are dark in color. However, if we focus on details, the movement of the birds during flight differs, the calls differ, and if we imagine them up close we would see significant differences in plumage. So in the first case the birds seemed very similar, but they seem very different in the second. Since, according to SAM, we ignore what is different between the two birds the level of similarity depends upon the initial representation of the anchor. SAM does not point to a mechanism which determines the level of detail in our representation of the anchor. It does not seem easy, then, to move from a set of initial conditions (target and anchor representations) to the target being evaluated as larger than in the absence of the anchor. Proponents of SAM need not be responsible for explaining the process which determines the level of representation, but until we know more about that process we do not know if SAM can explain anchoring effects.

Additionally, further specification of the automatic process which searches for similarities by psychology may spell trouble for SAM with regard to desideratum (2). As it stands, SAM gives some explanation of anchor persistence. It is vulnerable, however, because further specification of the automatic processes involved in anchoring could show that process to be a 3rd party cause. That means that anchoring effects and semantic similarity could be caused by changes in the underlying mechanism separately. Again, until we know more about the underlying process involved we should at least be worried about the possibility. That worry should be especially salient given Oppenheimer et al's

results which show semantic priming in cross-modal anchoring. In that study, (seemingly) non-semantically related targets/anchors still bring about anchoring effects and semantic priming.

Finally, we might question just what is meant when SAM states that the more semantically related two objects are determines how close our target estimate will be to the anchor. This question relates to desideratum (1) but is especially pertinent to desideratum (3). Phenomenologically, it seems clear that a human and a chimpanzee seem somewhat similar, but an automobile and an asteroid are not similar at all. What isn't clear is how similar a cow is to a hippopotamus (or, say, an elephant). If the theory is going to provide some indication of the range of values which will 'feel right' for a target estimation, it seems like we need to be able to predict the level of similarity prior to evaluations taking place. Moreover, if we only pay attention to the ways in which the target/anchor are similar, what is the threshold of similarity needed for us to reject the anchor value and what determines which way to adjust from that value? To answer both questions it seems like we need *something* else to be part of the anchoring mechanism than merely similarity and adjustment. Needing that *something else* is the largest problem with relying on similarity to explain anchoring. Similarity is not specific enough for us to generate anchoring phenomenon and it merely gestures at a mental process instead of explicitly laying out the mechanism involved in determining how we decide what is similar and what is not. Not satisfying desideratum (1) has further implications for how we can solve the individuation problem which I will discuss shortly.

While the picture of anchoring that SAM provides is not yet sufficiently filled in to give us a full explanation of anchoring, there are a few key benefits to the theory which I want to highlight. First, SAM incorporates semantic priming which seems to occur in many cases of anchoring even when it is not clear that the target/anchor are semantically related. Second, it hints at a reason why target estimations 'feel right' (amount of similarity) even though a closer inspection leads to some troubling questions. Finally, SAM's focus on similarity during the construction of our representations of entities

means that, relative to non-anchored situations, our target representations are affected by the anchor value. At a very basic level, then, it accepts that representations change and hints at a way in which those representations are altered (biased search for similarity). Focusing on how we construct representations can be beneficial when trying to explain anchoring. It should be noted, however, that this final point is contentious to some in the field.

Scale Distortion

Frederick and Mochon (2011) have put forward the Scale Distortion (SD) model of anchoring which rejects the idea that our representations of objects change during anchoring. SD says that we map our representation of the anchor onto an 'objective' response scale. The term 'objective' is merely meant to denote scales which are communicable (e.g., inches, miles, or pounds). When we use that scale to evaluate the target, however, our representation of the target feels 'too large/small' by comparison. We then "shift the response scale" (1) until it matches with our representational scale.

The easiest way to understand SD is by the example that Frederick and Mochon provide. Imagine that in the first phase of anchoring you are asked if a raccoon is smaller or larger than 20lbs. You then map your representation of the raccoon to the 'pounds' scale to answer the question. In the second phase, you are asked to estimate the weight of a giraffe and you compare the representation of the giraffe to the same response scale as the raccoon. Let's imagine that, in a non-anchored situation, the giraffe representation would correlate to 1,000lbs on the response scale. 1,000lbs "feels" (2) too large in comparison to the twenty pound raccoon found in the anchoring question. To compensate, you shift your response scale so that the giraffe representation feels more appropriate to that scale. The result is that the giraffe representation is lower on the 'pound' scale (e.g. 700lbs). The shift in the response scale, then, accounts for anchoring effects without having to say that our representations of the giraffe or

raccoon were altered in any way. Frederick and Mochon's experiment described in section 2 was taken as evidence in support of their model.

When it comes to satisfaction of the desiderata in section 3, SD is going to have to rely upon the content of our mental representations to provide an explanation. Information about the comparative size of the animals, as well as information about how to map that size onto a response scale all requires that we can get that information from our representations. The reliance upon representation is both a major benefit and a significant hindrance. Pointing to representational content is a benefit because it allows us to (partially) explain desiderata (2) and (3). It is a hindrance because SD mandates that the representation of both the animals and response scale be unchanging.

For desideratum (2), SD can say that the anchor information persists during the target evaluation because we use the same distorted response scale to generate our answer. Anchoring persists even when you inform subjects about it because subjects cannot ignore the feeling that the target is either too small or too large on the response scale. The representation of the target provides us, presumably, with all the information that is necessary to reject the anchor, give us a sense of the direction we need to adjust, and provide some range of values which feel reasonable for the target. Therefore, we have all the basic ingredients for an explanation that captures desideratum (3). However, the reliance upon completely static representations would seem to cause a problem for the theory. If all representations are static, then whatever scale we use should keep representational information perfectly proportional. If our representation of raccoons and giraffes includes static weight information, then we should be able to express giraffes in multiples of raccoons. Once we map the representation of the raccoon to the response scale, the giraffe representations should be perfectly scaled. If everything scales perfectly, we should *not* get the 'feeling' that the giraffe representation is too big on the 'pound' scale. Therefore, in order for SD to keep the theoretical claim that representations are static, SD will require some other component to tell us why the giraffe feels too big. If SD needs an addition to derive

the phenomenon, then that should tell us that the theory does not satisfy desideratum (1). Proponents

of SD might amend the theory to say that representations are not be static, but still insist anchoring is

caused by shifts in the response scale. Such a tactic would still not allow them to satisfy the desideratum

because it misses a more fundamental issue. The biggest issue for SD is that it points to the feeling of the

giraffe being too big as opposed to the process or mechanism which is responsible for that feeling. If we

have the feeling the giraffe is too big, then we have to get that information from something but SD does

not tell us what that *something* is.

I will not fully examine the explanation of anchoring that Oppenheimer et al. put forth

because they do not consider Magnitude priming to be a "rival hypothesis to the other major theories of

anchoring" (Oppenheimer, LeBoeuf, and Brewer 2008, 24). However, it is worth sketching out the theory.

Magnitude priming roughly is the idea that anchors prime us for general *largeness* and *smallness.* When

the anchor is large, we are primed to think that large values are more appropriate. The thing I find

interesting about Magnitude priming is that for something to be *large* or *small* means our

representations have spatial information encoded in them. Without spatial information being part of the

representations, we would not be able to categorize objects as being small or large in the first place. If

something like magnitude theory is right (at least in conjunction with other theories), then it implies that

representations are, in some ways, susceptible to changes in the scale of that spatial information.

There has been a lot of information in this section so I have included a table of the various

theories, their advocates, and a brief description of the theory below.

Model		Explanation	Desideratum		
			(1)	(2)	(3)
Anchor and Adjust					
Tversky and Kahneman		Result of adjustment process based upon the initially presented value. Adjustments continue until value seems reasonable.	Underspecified	No explanation provided	No explanation of range or direction of adjustment. Explanation of stopping point is phenomenological.
Eply and Gilovich		Anchor provides a reference point which defines the range of plausible target values.	Underspecified	No explanation provided	Range and direction of adjustment provided by anchor. Stop occurs when value falls inside the accepted range.
Selective Accessibility (SAM)					
Mussweiler and Strack, Chapman and Johnson		Anchor value is tested as hypothesis and adjustment amount is the result of a biased search for semantic similarity between the anchor and target	Underspecified	Persistence result of activation residue.	Range, direction and stopping point all defined by level of 'similarity'.
Scale Distortion (SD)					
Frederick and Mochon		Initial target value seems too large/small, relative to the anchor value, on an objective scale. We adjust target value until proportionality 'feels' correct.	Underspecified	Persistence based upon shifts in response scale	Range and direction based upon reference to shifting scale. Stopping point is phenomenological.

In discussing how the various theories addressed the desiderata from section 2, I deliberately avoided discussing desideratum (4). As I said in that section it would be nice to have a single theory account for every instance of anchoring, but whether all anchoring is caused by the same mental process is an open question. SAM is limited to those examples where the anchor and target are semantically similar. What, exactly, semantic similarity means is not clear and so there is no way to judge (at the moment) how large the range of anchoring effects SAM can accommodate.

As it stands, however, SAM is unable to account for either Oppenheimer et al or Frederic and Mochon's studies. It does not seem to explain Oppenheimer et al's study because it is hard to see the semantic link between a line drawn on a piece of paper and the Mississippi River. Perhaps the response would be that the Mississippi River is conceived as a line on a map and so the concept *line* is what is similar. If what counts as similarity is so broad, then we should wonder why more of our judgments are not significantly affected by everyday objects. Even when semantic similarity is easier to find, SAM is unable to tell us much about the anchoring study by Frederick and Mochon. SAM does not tell us much about this study because it does not provide a way to figure out whether a wolf is more similar to a cow, giraffe, or hippopotamus (to use random examples).

The way SD is currently structured it does not fare much better than SAM. SD cannot explain the anchoring examples of Jacowitz and Kahneman or Mussweiler and Strack because the anchor object and target object are the same entity. It is not clear from the theory how a shift in response scale is possible when you are evaluating the same object (like the Mississippi River) in both stages of the experiment. It appears that SD only works when the target and anchor are different entities. Nor can SD easily explain Oppenheimer et al or LeBoeuf and Shafir. In LeBoeuf and Shafir, participants were allowed to examine a 3.5 inch line in one room before reproducing it in another. Since representations in SD are static, subjects should be able to recall the original representation for comparison. SD does not seem to explain Oppenheimer et al because any line you draw on a piece of paper is tiny compared to the length of the Mississippi River. SD would predict that both the "short" and "long" anchors would make the Mississippi River feel 'too big' and, therefore, be small anchors. Perhaps there is a reason why drawing a 'long' line on a piece of paper acts as a large anchor for the Mississippi River that SD could accept. At the moment, however, that reason is missing from the SD model. Both SAM and SD, then, suffer from being underspecified. Underspecification does not mean we should automatically reject them, it just means that there will be difficulties in addressing desiderata (2), (3), and now (4).

It could be argued that the examples I have provided in section 1 are different from one another. Even though they all are examples of anchoring effects, they could be the result of several different mental processes. In such a case, we should not expect the theories to be able to explain each of the examples. Subjectively, it does seem like I *do* different things in each example (e.g. drawing lines, comparing animals, estimating rivers, etc). Assuming that the examples I provided are served by different mental processes, then we should be asking how can we know which type of anchoring process will be employed in any given anchoring study. In other words, how does each theory address the *individuation problem*?

To my mind, desideratum (2) and (3) are the most critical to addressing the *individuation problem*. All the theories we have looked at merely gesture at the phenomenology involved in desideratum (3). Gesturing at phenomenology is insufficient for either individuation or explanation, regardless of how accurate the description is. What we really want is a mechanism which explains both why target and anchor values 'feel right' (or wrong), and is known to resist alteration by higher level processes (desideratum 2). Specifying the different mechanisms involved in anchoring are what, ultimately, solves the *individuation problem.* When the mechanisms involved in one set of anchoring phenomena differ from the mechanisms involved in another set, we know that the two groups of phenomena should be considered distinct.

Because the three major theories we have looked at do not address the underlying mechanisms involved in desiderata (2) and (3), they are unable to individuate anchoring phenomena. To my mind, the problem is that the theories have started from the phenomenon and tried to work backwards to the mechanisms. If we want to specify mechanisms which both satisfy our desiderata and solve the individuation problem, perhaps, we should start with low level mechanisms to find out if they could produce anchoring effects. In the next section I will argue for focusing on the scale of our representational space as a candidate mechanism.

5   The Sticky Space Model

Imagine the task of throwing a ball of paper into a trash can that is four feet away. When you think about it, the total number of variables that need to be represented to accomplish the task is enormous. We need to consider the size and shape of the opening, the air currents around us, the angle the trajectory will take, as well as the size of the ball. Next, we also need to represent the force and speed of our arm as it travels forward and the point where the ball leaves our hand. Then we need to represent the ball as it moves away from our position towards the hole in a manner that is consistent with all the forces that will act upon it. Assuming that we were able to make our 'shot' we may even represent the sound the paper ball makes as it drops into the can. That sound, which has to arrive at just the right time, has its own properties that have to be adjusted relative to our normal hearing ability, the sound things of that size make at that range, and other conditions like wind or background noise. Moreover, we have to have representational content not only about 4' distances but also 1', 6', or 40' throws. We also need information about those throws when the can is uphill or downhill from our location. If all of that information was available simultaneously, some of the information we have about the task will conflict with other information. For example, the information about how hard to throw the paper ball at a target 4' away is much different when the target if 20' away. Having to sort through all the information to choose which information is useful in our particular context would take an incredibly long time. Yet, we are able to accomplish the task fairly quickly. Therefore, it seems more reasonable to think that our representations are not nearly as complete as the above picture suggests. The more economical approach to the problem is to only provide context relevant information.

If we want to deny that our representations are 'complete' in the sense above, then we will need some way to access only the information that we need. Some of the information listed above is specifically necessary for the task of throwing the paper ball, others for predicting the path of that ball, and still others for tertiary information like the sounds, smells, etc. How do we to pick out the stuff we

need while ignoring the stuff we do not? The question should sound familiar because we encountered a similar question when examining SAM. For the current task it will be helpful to break the question down into more basic questions. We need to ask how it is that we represent distance, movement, and speed. We also need to ask how visual, auditory, and tactile information gets depicted and when. We might further ask how our motor functions get incorporated into the task, but this question will not be particularly relevant to the rest of the paper and so I will omit it.

In our imagined trash can scenario, we need to represent the paper ball being X distance to the can. The first question I want to put on the table is how do we manage to fill in X? One way to go about the inquiry is to think about how we move and act in a given space. In the early 1970's it was discovered that there were cells in the hippocampal system of rats which seemed to activate in very specific locations (O'Keefe and Burgess, 1996; Hafting et al., 2005). Place cells, however, do not create a map in any ordinary sense of the word. While a cell might become active at a particular location, the neighboring cells did not become active when the rat moved away from that location. In other words, adjacent locations in the environment could activate cells that were far away from each other. To complicate matters, sometimes cells would 'remap' the location that they responded to.

More recently, biologists have discovered a group of cells, called 'grid cells', in the entorhinal cortex of rats. The entorhinal cortex is a portion of the brain which feeds information to the hippocampus; an area which has been implicated in spatial manipulation tasks (Clearwater and Bilkey, 2012; Zhang and Ekstrom, 2013). Recently, evidence for a functionally similar architecture has been found in humans (Doeller, Barry, and Burgess, 2010). Grid cells act *as if* there was a grid of equilateral triangles overlying on the environment. Objects and locations in a given space correspond to a particular intersection of the overlying grid. So, when a rat moves from point A to point B, it can keep track of where it is in relation to other objects by monitoring where those objects are located on the grid. The density of grid cells changes as you move from more dorsal sections of the brain to more ventral regions.

As the cell density changes, so does the size of the grid overlay (Brun et al., 2008). Since information about an objects location only is transmitted when the object occupies a vertex of the grid, the size of the grid will determine the amount of information that is available about a scene. In rats, the most dorsal cells represent a grid size of about 30cm (roughly the size of the rat), but at the most ventral regions that grid size is closer to 3m (Brun et al., 2008). Objects in the environment have a significant impact on where the activation happens in the brain, and therefore, how big or small the grid is (Hafting et al., 2005; Wang and Spelke, 2002). Interestingly, when you remove environmental stimulus, say by turning out the lights, the grid does not immediately disappear. Instead, the grid stays static which allows for the rat to navigate without constant updating from the visual system.

Grid cells, in conjunction with place cells and head- direction cells, form part of the neural infrastructure that allows for the allocentric representation of the environment. The allocentric frame represents objects from a view-independent perspective, much like a physical map represents objects (or locations) in a way that is indifferent to where *you* happen to be. We also have a view-dependent, or egocentric, representation of the world which locates objects in relation to our body. The egocentric frame is typically associated with the parietal cortex. This area is also implicated in the mental rotation of both real and imagined objects (Just et al., 2001). The egocentric and allocentric frames constitute parallel streams of information which can be disassociated (Burgess, 2007). Even though the streams are dissociable, there are two important things to keep in mind. First, the two parallel streams are typically integrated and both streams provide information which is used to navigate and act in the world (Burgess, 2007). Which information stream takes precedence during integration is partially a function of the information that is available in the environment as well as the context of the situation. For example, when allocentric cues are unavailable during a navigation task, people tend to rely more upon egocentric information to find their way around. Second, research has shown that allocentric information can be converted into the egocentric frame (Galletti et al., 1995). Some evidence shows that translation of the

allocentric representation is necessary to influence motor behavior (Byrne et al., 2010; Chen et al., 2011). The exact mechanism of translation is still under investigation, but the hippocampus and entorhinal areas of the brain have been implicated (Bird et al., 2012; Hafting, 2005).

Recently, scientists have begun to think that human navigation, prediction, and memory/imagery are all (partially) served by the same mental infrastructure (Hassabis and Maguire, 2007). Of particular interest for this paper is the idea that allocentric representations, which rely upon the geometric overlay of grid-cells, are involved in imagery and imagination. The allocentric system is involved in the construction of imagery in two ways. First, the allocentric system undergirds the spatial coherence of the imagined scene. For imagery to be spatially coherent means that the individual items of a scene are not represented as individual fragments. Rather, the entire scene is constructed holistically (Hassabis et al., 2007). The evidence that the allocentric system is involved in spatial coherence is found in lesion studies by Hassabis et al. (2007). The researchers found that lesions in the hippocampal system impair a patient's ability to imagine scene components as part of an interconnected 'picture'. The ability to make a single coherent scene during an imagery task is important to keep in mind because the ability requires us to proportion spatial relationships between elements. In everyday behavior the ability to keep spatial relationships proportional is necessary for generating useful outcomes. If the same ability transfers to imagery, then we should think that the spatial relationships in imagery are also (generally) accurate enough for behavior and prediction. More will be said about this point later in the paper.

Second, the allocentric system is partially responsible for determining what kind of non-spatial information is recalled during imagery. Bird et al argue that the allocentric system provides a "coherent spatial context within which imagery can take place."(Bird et al., 2012, 10) When the spatial context provided by the allocentric system was missing (or weak), Bird et al found that egocentric information can dominate our imagery when that information is strongly encoded.

What is important, for our purposes, is that our representational system is made up of several lower level processes. Each process provides information about our world and our relationship with it. The integration of those processes appears to serve both real-world behavior and our imaginative abilities. The ability to keep track of objects in the environment relative to their location on a spatial grid allows for the common (human) ability to interact with objects that are outside of our field of view. To see what I mean, place a pencil (or some other object) somewhere within arm's reach. While keeping your eyes focused on this sentence, reach out and pick the pencil back up. To accomplish the task you had to use information about where the pencil was relative to you and other objects nearby (e.g. your computer mouse or keyboard) so that you reached in the right direction, to the right distance, and grasped the pencil instead of something else. All of that was done without having to obtain the information visually.

To pick up the pencil, the size of our spatial grid needs to be fairly stable. Since objects in the environment determine the size of the spatial grid, we can infer that our environment was also fairly stable. The initial question we should be asking ourselves then is "What kinds of things do we tether the spatial grid to?" When walking around a city our grid might be set by the skyscrapers, trees and mountains that make up stable aspects of our visual scene. However, it is not just skyscrapers, trees, and mountains that tether our grid to the environment. Objects that define the space around them appear to also include things like couches and bookcases (Mullally and Maguire, 2013) and walls (O'Keefe and Burgess, 1996). Having a wide range of objects which we can tether our grid to is important. The more objects to which we can tether our grid means the more adaptable our grid will be. Having an adaptable grid is really useful. If our grid could only be tethered to mountains and trees, for example, we would be unable to change the scale of the grid when we entered buildings. Since information about an object's location is only available when it occupies a vertex of the grid, the large scale grid set by mountains and trees would make it impossible to track items in the (relatively) small spaces of our homes and offices.

So, we want to have a system that can adapt to the variety of locations which we commonly find ourselves. We don't want, however, a grid that changes whenever objects in the visual scene change. For example, we do not want the representational scale of the room to change any time we look out the window at mountains and trees. Fortunately, the fact that we can pick up pencils without looking at them suggests that the grid *doesn't* change immediately regardless of whether we want it to or not. We are now in a position to say when our representational space is 'sticky':

> *Representational space is 'sticky' when the scale of the representational space tends to persist despite changes in environmental stimulus.*

Having a 'sticky' representational space is quite beneficial. It allows us to shut doors, pick up objects, and track the location of items when they are not in our field of view. It also helps to preserve information that is highly relevant to our actions even when there are alterations in the visual stimulus. Moreover, the fact that we are able to do all of these things without having to consciously think about it tells us that the system is fairly reliable. If the system was not generally reliable, actions that depend on utilizing our spatial grid would not typically have the consequences that we want.

Not only is the scale of the spatial grid useful for navigation and action, but there is some evidence that suggests the scale of the spatial grid fundamentally changes the way we represent objects in our environment. For example, a study by Meyer-Levy and Zhu (2007) shows that the scale of a room can affect how we describe and categorize objects in that room. The researchers manipulated perceptions of ceiling height by hanging paper lanterns either a) very low or b) very high. When study participants viewed a scene where the lanterns were hung very low they described objects in concrete terms and categorized objects into many different groups. When study participants viewed a scene where lanterns were hung high in the ceiling, they used fewer categories and described objects by

focusing on abstract features.[2] If object descriptions and categories tell us anything about what representational content is salient to the study participants, then it appears that what content are salient changes with the scale of the spatial grid.

As we have already seen, objects in our environment are what help to set the scale used to navigate that environment. It should not be controversial to think, then, that a high ceiling is going to create a scale which is larger than a low ceiling. Typically, when provided with a small (or 'zoomed-in') scale, certain representational content becomes more salient and important. Conversely, that information may not be salient under larger scales ('zoomed-out'). If the scale of our representational space is large, we will typically include only those features which would be appropriate for that scale. Since there is typically less detailed information available when we 'zoom out' from objects, it should be no surprise that descriptions of objects in high ceiling scenarios are less concrete.

Additionally, since the level of detail in the representational content increases as the scale decreases, smaller scales increase the distinctiveness of objects in the scene. In other words, the more we 'zoom-in', the more likely we are to have the differences between objects made salient. A focus on the differences should lead us to see less connection between the objects. Therefore, we are less likely to categorize the objects together. For example, from far away a tennis ball and a racquet ball are very similar. They may differ on color but the relative size and shape is the same. A closer inspection reveals the differences in physical properties which affect how they interact with the world. Similarly,

---

[2] This study is part of a larger field of research into the effects of psychological distance as put forward by Construal Level Theory. Construal Level Theory claims that psychological distance (either spatially, temporally, or socially) changes the level of abstraction with which we construe situations which then effects behavior (for overview see Trope, Liberman, Wakslak; Construal Levels and Psychological Distance: Effects on Representation, Prediction, Evaluation, and Behavior; *Journal of Consumer Psychology*; 2007;17(2): 83-95). I do not aim to explain psychological distance nor does that research explain anchoring effects. The use of the data here is only to suggest that it is plausible for spatial scales to influence what representational content is made salient about an object. This becomes important to anchoring when we consider one of the problems with SAM was underspecification about what content we find similar. More will be said about that later in this section.

when our representational space is 'zoomed-out' we find the two objects generally the same and would group them together. When representational space is 'zoomed-in' we notice the important differences and may be more likely to assign them to different categories.

If you recall from our earlier discussion of SAM one of the limitations of that theory was that it was hard to tell what level of abstraction we employed in determine what counts as a semantic similarity. The example I gave there was that of a crow and a starling. We are now in a position to be able to fill out why the two might be considered similar in some situations but dissimilar in others. When we compare the birds, the information about their plumage is determined by the spatial scale of our representational space. If our representation of the birds is zoomed-in, then the spotty and iridescent plumage of the starling contrasts with the flat black plumage of the crow. If our representation of the birds is zoomed-out, then both birds will have 'dark' plumage. If a crow is used as an anchor for size estimates of a starling, whether the anchor is effective will be a direct result of which representational scale is used. In theory, the more abstract the objects are perceived to be, the more likely they will be seen as similar. So, if SAM is correct, greater anchoring effects should be seen when items are perceived as more abstract.

Before moving on, I wish to bring up one point in regards to SD. The discussion of how the scale of our spatial representation system affects the content of our representations, in some ways, speaks against Frederic and Mochon. If you recall, they claim that all our representations are static entities. In some sense that may be true. If we consider all of the information that we could possible represent about an object, that information may be unchanging. The content that is salient, and therefore open to manipulation, however, does seem to vary.

To recap this section, I have attempted to show three things. First, that we have a neurological basis for accepting the idea that spatial representation relies upon the scale of a spatial grid. Under normal situations, like walking around the city, that grid stays relatively stable in size. The scale of

our spatial grid, however, is not entirely static. External features of the world do impact the size of the grid, and therefore impact what information is available to us. There are benefits to both a static and dynamic system and that has led to, what I term, a 'sticky' representational space. The scale of our representational space tends to persist despite changes in the space defining stimulus. Second, I have attempted to show that information we generate from having a sticky representational space is reliable and having a sticky space is beneficial. Having a sticky spatial grid is beneficial because it supports many of the day to day actions in which we engage. Moreover, because we successfully perform actions which rely upon a sticky spatial grid without having to consciously think about them, we can infer that the information the representational system provides is reliable.

Second, as the scale of the spatial grid changes, certain information becomes more, or less, salient and useful. Representations of objects which are placed in a large scale space tend to make certain features salient, whereas other features become salient in smaller scaled spaces. Therefore, the scale of our spatial grid has profound effects on how we conceptualize the world. If you recall SAM suffered from being unable to specify why objects are considered semantically related. The way our spatial grid affects what is salient in our representations, takes an initial step to filling SAM in. However, if the spatial grid is the mechanism by which similarities are made salient, and the spatial grid also brings about anchoring, then SAM has not carved nature at its joints.

So far though this conversation has been limited to how our representational space is affected by real-world objects in the environment. When we ask our representational system to deal with counter factual, hypothetical, or imagined situations does this picture still hold? I believe it does but with some surprising results. In the next section I will argue that when we use the scale of our spatial grid to gather information about objects in our representational space, without the benefit of the real-world tethers, we are at risk for introducing distortions into the information. I then turn to examining an anchoring example to show that the distortions can generate anchoring effects.

6   Sticky Space and Anchoring

Imagine a simple house cat. Everyone will probably fill in many of the details about that cat differently. Whatever size and distance you imagined the cat is not particularly important (yet), just notice that the representation of the cat was at *some* distance and visual size. With the cat firmly in mind, estimate how tall the cat is in feet. If you happen to know, as a piece of trivia knowledge perhaps, how tall house cats generally are then you could recall that value from memory. Otherwise, you will have to rely upon the information found in your representational space. Using the representational space forces you to represent the cat at a scale where such information would be available to you. What I mean is that if your representation of the cat was very abstract (e.g., you used discursive imagination with no imagery at all) or you represented the cat as being very far away, you will not have the relevant information in the representation because there are two few grid points to map onto the cat. As a result, you will have to 'zoom-in' on the spatial grid until you have enough points to make the determination.

Let's say you now have the cat represented in a manner which the height of the cat in feet can be determined. Now imagine there is a full size bull elephant standing right next to the cat. If you are like most of the people I (unscientifically) polled one of two things happened. Either you imagined the entire elephant and the size of the cat shrank, or the visual size of the cat stayed the same but you only imagined the elephant's leg. Now I want you to estimate how tall, in feet, the elephant is.

If we were actually seeing an elephant next to a house cat we could use the visual information available to us to help answer the question. Whether imagination uses a pictorial format is a matter of some debate in philosophy. There are those that would argue our representations are depicted quasi-pictorially (Kosslyn, Ball, and Reiser, 1978). Others argue that representations are propositional statements in the language of thought and the 'visual' experience is entirely epiphenomenal (Pylyshyn,

1973). How information is stored in memory is not important to this discussion. What is useful is the neuroscientific evidence which shows that imagery is processed on much of the same infrastructure as perceptual stimulus from the 'real-world' (see (Kosslyn, Ganis, and Thompson, 2001)). Even if we use the same infrastructure in imagination, imagined scenarios do not necessarily enjoy the stability of the spatial grid that we might otherwise find. The problem, for our purposes, is that it seems our representations of the cat and elephant want to use two different scales, but we need a single scale for the estimation task.

If your visual representation of the cat 'shrank' with the addition of the elephant, then you do not have enough grid points to provide information about height. Since we already estimated the height of the cat, we need cat-information to compare to the elephant representation. If your visual representation of the cat stayed static (and only part of the elephant was represented), then the entire elephant is not mapped to the spatial grid. In either case, the size the grid (or the size of the representations on that grid) needs to be altered in order to accomplish the task. If we want the entire elephant to be mapped onto the grid we are going to have to distort the size of the cat to do make our comparison. If we want to be faithful to the size of the cat, then we will have to add to the representation of the elephant in chunks, in order to keep the cat-scale constant.

Doing either manipulation is going to cause trouble. If we manipulate the size of the cat, then the reference scale of our estimation will be distorted. If we try to represent the elephant bit by bit, then we also have to extrapolate the information about height from the current information about the leg. Meaning we have to already have that information available if we want to accomplish the task with any precision. In either case, it seems like the task is going to be difficult to accomplish without any distortions creeping in.

We might wonder why these tasks are difficult. It seems like we could scale up or down with perfect proportionality, in which case, the comparison between the cat and elephant should elicit perfect scaled estimates. I want to resist that idea because the ability to manipulate and scale objects in the representational space with high degrees of precision does not seem to be something most people enjoy. Consider that there is a high degree of variation in people's ability to perform mental rotation tasks. Some people are able to do the tasks quickly and accurately, while others cannot. Mental rotation tasks require us to manipulate and scale objects in our representational space. So, it is unlikely that everyone enjoys the same ability to manipulate and scale representations more generally. The distortion of objects due to imperfect scaling and manipulation are a path, I will argue, to response values that exhibit anchoring effects.

We are now in a position to apply the Sticky Space model to anchoring phenomena. As you may have noticed, the previous example in the last section was modeled loosely on the animal comparison study done by Frederick and Mochon. We can take a step further and apply the model directly to experiment. In that experiment, subjects were asked to first estimate the weight of a wolf in pounds. Next, subjects were asked to choose the animal (from an ordered list) that they thought was closest to 1,000 pounds. In comparison to those who did not have to estimate the weight of a wolf, the test subjects chose animals which were significantly heavier during the final task. The Sticky Space model also predicts this to be the case. During the anchoring phase of the experiment, subjects retrieve spatial information about wolves and use that information to map the representation to an objective value in pounds. The representations of the animals which are considered response candidates are then integrated into the representational space. Since representations retrieved from memory come with different spatial scales, the integration of the new representations causes a fluctuation in the scale of the representational space. These fluctuations lead to distortions which affect the 'measurement' of the new

animal representations. Therefore, the representations of the new animals tend toward the weight of the wolf anchor and 'larger' animals (compared to the control condition) are finally selected.

A Sticky Space model, therefore, is not incompatible with SD. In broad strokes, in fact, it is quite similar. The difference between SD and the Sticky Space model is that the spatial information in our representations can be altered in the Sticky Space model. This small difference allows for us to explain why our representations do not scale proportionately.[3] Notice, however, that the Sticky Space model is not limited to comparisons of different objects like SD. It also predicts anchoring effects when the target and anchor can be the same entity. To show how this is the case, I would like to use the Mississippi River question from Jacowitz and Kahneman's study. I do this because the process will be very similar and using another studied case will give a sense of how many example of spatial anchoring a sticky representational space can accommodate.

If you recall, subjects were asked to first evaluate whether the Mississippi river was longer or shorter than an anchor value. For this example, we can use the short anchor value of 70 miles. To accomplish the task I first need to have some idea of what '70 miles' is. I know that the distance between Atlanta, GA and Athens, GA is roughly 70 miles, so I can use the representation of the distance between Atlanta and Athens (hereafter A-A) as a way to measure distances. The A-A distance is too small to accommodate my representation of the Mississippi River because that representation wants to be scaled to the size of the U.S. So, to 'measure' the river I need to zoom-out by expanding my spatial grid. That I have to zoom out already tells me that '70 miles' is too short of an estimate for the length of the Mississippi River.

In the next step subjects were asked to estimate the length of the Mississippi River in miles. Up to this point, I have not needed to do a comparison because merely zooming out gave me the

---

[3] Recall that one problem with a completely static representation should lead to (near) perfect scaling and thus not "feel wrong".

information needed. Now, however, I need to measure the length of the river. The only spatial

information that maps onto the appropriate response scale is contained in the A-A representation. So my

representation of the Mississippi River has to be compared to the A-A representation. My representation

of the Mississippi River is in the context of a map of the U.S., but the A-A information is too small to map

onto multiple grid points at that scale. To proceed with the task, the A-A representation is adjusted so

that it can occupy at least two grid points. Since the scale of my spatial grid is sticky, and I am unable to

perfectly scale the representation, that representation is slightly larger than it would otherwise be.

Therefore, the spatial information which I equate to 70 miles is no longer as accurate as it was. When I

then read off the length of the Mississippi River based upon my spatial grid, it is based on a scale which is

too large. Meaning, that my absolute estimate will be smaller than it might otherwise be. Because I am

able to read the length of the Mississippi River off of the spatial grid, I get the subjective feeling that my

(small) response value seems reasonable. If the response value coheres with my background information

regarding how long rivers can be, then I report the response value to the researcher and my task is

complete.

Finally, because the scale of our spatial grid helps determine the content of my

representations, and my grid is biased toward the anchor, it is plausible to assume that I will be primed

for anchor consistent semantic content. Obviously, this is an empirical claim which should be tested, but

I would think that in this situation I might be primed for generally abstract words and descriptions

because of the relatively large scale of the anchor.

The studies by Frederick and Mochon and Jacowitz and Kahneman are instances where

both the target and anchor have to be imagined in order to accomplish the tasks.[4] What about

---

[4] It might be claimed that Frederick and Mochon's study does not require the use of imagination in the final phase
because participants use the visual size of the animal pictures found in the evaluation phase. I am dubious of
the claim, but there is a quick way to test it. Repeat the anchoring study using 'stretched' images. Stretched
images should appear visually longer and skinnier and, therefore, less massive. If subjects preference the

anchoring studies which force us to compare recalled information to immediate stimulus? The study by LeBoeuf and Shafir is a prime example of such circumstance. In that study participants viewed a 3.5 inch line in one room and then needed to erase or extend a line in a separate room. Participants who had to erase a long line tended to overestimate the length of 3.5 inches. Participants who extended a short line underestimated the length. The Sticky Space model is able to accommodate these results. During the initial examination phase of the experiment, subjects encode the spatial aspects of the line relative to the scaling of their representational space. That scale will be set by various features of the room as well as the relationship of the line to those objects. To get to the test location, participants had to navigate to another room. It is entirely open for the scale of participant's representational space to be altered by the differences in spatial features found in the test room. Those features can include the relationship of the target line to surrounding objects or boundaries. When the participants recalled the original line there would be a conflict in scales. That conflict could precipitate an integration phase which results in the distortion of the *recalled* line. Therefore, the recalled line would be biased towards the immediate scaling, which includes the target line. Obviously there are going to be limitations on how quickly the scale of our representational space changes. Those boundaries are an empirical matter and if empirical data shows that moving rooms in unlikely to cause a shift in the scale of the spatial grid, then that evidence would count against the Sticky Space model. Whether the Sticky Space model ultimately explains LeBoeuf and Shafir's results is (at present) an open question, but there is no immediate reason to think the results are contradictory.

The Sticky Space model is unable however, to explain the cross-modal results of Oppenheimer et al. without having an additional component. I will briefly discuss how the theory might be extended to cover cross-modal anchoring at the end of this paper. As it stands, the Sticky Space model fares no better

---

immediate visual stimulus over mental representations for the task, then they should pick animals higher up on the scale. If there is no statistical difference in responses, then it implies subject's preference their internal representations.

than any of the other theories we have examined. With that said, there are a few advantages to the model which should count in its favor.

First, and foremost, the Sticky Space model is able to capture desiderata (2) and (3). Anchors along spatial dimensions are persistent because we leverage the scale of our spatial grid to generate answers for questions about which we are uncertain. The scale of that spatial grid is sticky, meaning that it tends to persist even when the environmental stimulus changes. When we perform the evaluative task at of the traditional anchoring paradigm at t_1, the scale of our representational space gets untethered from the environment. The mental representation of the anchor object at t_2 sets an initial scale for the representational space. Since objects in our representational space are, partially, responsible for the scale of the space we cannot completely bypass the affects of those objects via top-down strategies.

The Sticky Space model also provides us with a reason for why we reject anchor values, have some idea of the direction in which we need to adjust from, and why our final estimations 'feel right'. Continuing along the time line we began in the last paragraph shows us how we generate the values. At time t_2, the scale of the representational space was initially set by the anchor. During the estimation task we add the target into the representational space (t_3). The representations of the target and anchor want to set scales that compete with one another at t_4, and that competition tells us the values are not identical. The difference in the scale the target attempts to set tells us which way we need to adjust from the anchor. At t_5 the scale of the representational space becomes more 'fixed' and we are able to use that scale to arrive at the estimate of the target. The estimate 'feels' right because we utilize the scale of the spatial grid to generate the value and that process typically gives us accurate information about objects in the representational space. The spatial information we generate, however, has been distorted during the scale setting process and so the values we arrive at are biased towards the anchor.

Finally, by specifying the mechanisms involved in (2) and (3) the model is able to derive anchoring effects in certain spatial domains (desideratum (1)). Unlike SAM, the Sticky Space model it is not limited to examples with semantic similarity; even though it is able to incorporate semantic similarity into the model. The difference between SAM and the Sticky Space model is that the sticky space model holds that the scale of our spatial grid causes both the anchoring effect and the semantic similarity. Unlike SD, the Sticky Space model is not limited to a particular evaluative dimension (e.g. length, width, etc). At the moment, however, it does not allow us to derive anchoring effects in other domains (such as numeric anchoring effects). With that said, there may be a way to extend the theory in order to accommodate a larger array of examples. I will briefly discuss how that might occur in the conclusion. Before moving on to the conclusion, however, I want to discuss how spatial anchoring is situated in the larger Bias and Heuristic literature. In the next section I will focus on the impact spatial anchoring (as I describe it) has on questions of Rationality.

7   Spatial Anchoring and the 'Rationality Wars'

Anchoring, as a research field, originally was part of a larger investigation into the ways human beings deviate from the norms of rationality associated with formal logic and probability calculus. The investigation has been headed up by advocates for what has been labeled the Heuristics and Bias program (HB). Advocates for HB, like Kahneman and Tversky, called into question whether people's judgments and decisions could be considered *rational* since the experimental evidence they gathered seemed to suggest that humans do not employ the rules of rationality. The experimental evidence cited by proponents of HB sparked off what has since been called the 'Rationality Wars'.

On one side of the fight were psychologists like Kahneman and Tversky who argued that anchoring effects, the conjunction fallacy, and base-rate neglect show that people tend to use simple rules-of-thumb (heuristics) to form their judgments. These heuristics, many times, arrive at correct judgments, but are subject to "systematic errors" (Kahneman and Tversky, 1996). It is important to note that HB does not say we are *incapable* of using anything but simple heuristics. Nor does HB claim that we *never* use unproblematic process to form judgments.[5] Rather, they claimed, we *tend* to use simple heuristics in order to save cognitive resources. The important thing to keep in mind, for the discussion on anchoring, is that it remains open for humans to use more resource intensive processes which would not be normatively problematic.

On the other side of the fight were proponents of the Bounded Rationality (BR) model. Proponents of BR claimed that our judgments and decisions were formed by 'modules' which evolved to deal with certain problems faced by our hunter-gatherer ancestors. These modules were highly adept at processing information so long as the right types of inputs were used. Proponents of this view argued that our ancestors did not have access to the language and rules of formal probability but they did have a module which processed frequency. When we put probabilities into the language of frequencies, errors

---

[5] I am using the most charitable interpretation of the Heuristic and Bias program as found in Samuels, Stitch, and Bishop (2000). Under that interpretation, much of the language which implies that HB proponents held the much stronger view, that humans can't use anything but "shoddy software" (to use Gigerenzer's terms), can be chalked up to mere rhetoric.

like the conjunction fallacy are made less frequently (Gigerenzer and Hoffrage, 1995; Cosmides and Tooby, 1996). Instead of interpreting the experimental evidence that Kahneman and Tversky gathered as evidence that we are terrible at probability, advocates for the BR held that the experimental results were merely a function of the format information was presented in.

Most of the details of the 'Rationality Wars' are not relevant to the discussion in this paper. There are two disagreements, however, which I think are important. First, there is a debate about whether HB needs to put forth an explanatory theory of the cognitive processes involved in reasoning. Second, there is an unresolved question about what we *should* accept as a norm of rationality.

The first disagreement stems from Gigerenzer's complaint that HB has not provided a falsifiable theory of the psychological processes involved in biases and heuristics. The problem, for Gigerenzer, is that the psychological description provided by HB is not explanatorily useful because it is woefully underspecified. Kahneman and Tversky use terms like anchoring and representativeness bias (to which we could add more recent terms like confirmation bias), but the actual processes these terms refer to are left too vague to count as explanations. Moreover, the vagueness of the descriptions do not allow for falsification because any experimental data where the response value is different than the expected value can be made to fit a bias or heuristic. As Gigerenzer (1996) says "The problem with these heuristics is that they at once explain too little and too much. Too little, because we do not know when these heuristics work and how; too much, because, post hoc, one of them can be fitted to almost any experimental result." (592)

In response to the claim that HB has not offered any substantive theory which would show how heuristics influence judgments and decisions, Kahneman and Tversky note that "Much good psychology would fail this criterion."(Kahneman and Tversky 1996, post script) Moreover, Kahneman and Tversky claim that mandating process models for heuristics is not a good idea. They point to the benefits of Gestalt theory as a prime example for both claims (ibid).

The second area of disagreement between the two factions of the Rationality Wars turns on differing conceptions of what we should adopt as a normative standard of rationality. While both

parties can agree that our norms of rationality need to accommodate specific outcomes (like not allowing for a conjunction fallacies), *how* we arrive at the outcome is equally important. What is at stake here is what processes we should take as a guide to rational decision making. Those in the BR camp take success as the metric of rationality. If a heuristic is successful at generating the values we want it to generate, then that process should be considered rational; even when it does not conform to the rules of formal logic or probability calculus. The focus on success is interesting because it shows how the rules of rationality (as used by HB) might be problematic For example Todd and Gigerenzer (2000) argue that complex statistical evaluations with lots of information tend to be worse at assessing heart attack risks than simple heuristics. If that is the case, then it would be irrational, on BR's use of the term, to follow the 'rules' that HB focuses on. Instead, we should jettison adherence to those rules because they are a poor guide to decision making in complex problems. If the proponents of HB are seen to be advocating for a certain conception of rationality as the proper guide to decision making, then proponents of HB should be worried about Gigerenzer's argument. They should be worried because they seem committed to the assumption that the picture they advocate for will always produce the best outcomes.

Proponents of HB, however, can respond to the guidelines concern by claiming that their investigation is only concerned with the internal validity of the standard picture. In other words, what they are most interested in is how humans deviate from an *idealized* agent. There are consequences for taking such a route which may not be amenable to extending the HB program. As Wallin (2013) notes: "[First] refusing to deal with decision outcomes should also lead to an abandonment of arguments referring to potential consequences of not following particular decision rules, such as money pump arguments. Second, [HB] guidelines will not be particularly convincing if they are completely dissociated from actual decision outcomes" (475)

The reason why I want to highlight these two areas of disagreement is because they directly affect much of what I have done in this paper. As might be expected, I am sympathetic to many of Gigerenzer's worries. Much of my critique of the major theories of anchoring center turns on the idea that the major theories are underspecified. All theories and models are, to some degree, underspecified

(including sticky space) so underspecification is not a problem full stop. The trick is to find the degree of specification which allows the theory to be explanatory. The problem I identified with describing the psychological processes involved in anchoring effects at too high a level, is that it allows for any instance where judgments cluster around a prior value to be labeled as irrational anchoring. Take, for example, SAM's focus on semantic similarity. It is entirely possible for a normatively unproblematic process to produce a judgment that tends toward an anchor value and have priming occur separately. Underspecification by SAM prevents us from knowing whether it is priming or an unproblematic process which resulted in the judgment. Moreover, it is commonly asserted that there are many (underspecified) processes which can cause anchoring effects. Because there has been no a priori way to establish when a particular theory will be involved in a specific instance, there is no way to ensure an experiment has controlled for the relevant variables. For example, because semantic similarity is vague, there is no way to rule out SAM as an explanation of Frederick and Mochon's animal comparison experiment. Even though it is not clear how a wolf might be similar to the various animals used in the experiment, it is still open for the connection to be made given the associative nature of our brains. Not only does underspecification lead to problems in the laboratory, but it leads to confusion when we try and understand how pervasive anchoring might be in everyday judgments because we cannot differentiate irrational from rational processes. In other words, almost any time a judgment tends toward a previously considered value, HB proponents can claim the judgment is (possibly) irrational.

I have argued that the degree of specification of the Sticky Space model allows for it to be explanatorily sufficient. Additionally, that specification allows for the theory to be falsifiable. If anchoring on spatial dimensions does not activate any of the areas used in spatial behaviors, then we can be fairly certain the theory is incorrect. Neither explanation nor falsifiability, however, tells us much about rationality. Now that we have some idea of what is involved in the rationality debate we are in a position to examine whether spatial anchoring is normatively problematic. Since the anchoring literature is part of the HB program, I need to compare a sticky space model to the standard picture. Since success

is, or should, also be important I will have to say whether or not the process should be seen as unsuccessful. I will address these issues in order.

Recall that according to the HB program we employ heuristics which violate the norms prescribed by logic and probability. Anchoring effects like those found in Tversky and Kahneman's U.N. participation experiment were considered irrational because participants used the available number (from the wheel) as opposed to generating a starting position to adjust from on their own. Similarly, under SAM anchoring is irrational because participants perform a biased search for similarity as opposed to considering both similarities and dissimilarities. In both cases, it is possible for participants to use a more effortful approach which would not violate a norm of rationality and, therefore, not generate an effect. We might formalize the rule that anchoring has been proposed to violate as: *Do not incorporate irrelevant information into judgments* (we will call this the relevance rule).

Does spatial anchoring, as described by the sticky space model violate the relevance rule? The first thing we should ask is if the anchor information is actually irrelevant in spatial anchoring. It is my contention that answering in the affirmative would be a mistake. The anchor value is relevant because it provides an initial mapping of our representational space to an objective quantitative value. Since our representational system attempts to make spatially coherent representations, meaning the spatial information of new objects must be integrated into the existing scale, the wolf representation has *informational relevance* for the integration task. Therefore, participants have not violated the relevance rule. Rather the final target value shows what is otherwise hidden: that our representations can be distorted. Since it does not violate the relevance rule, it should not be considered irrational on *that* account.

Perhaps spatial anchoring can still be said to violate something like the norm of transitivity. What is surprising about spatial anchoring is not that the number deviates from a factual value, but that the response we provide in an isolated situation changes during a comparison task. For example, it is not surprising that people do not provide accurate estimates of the Mississippi River. It *is* surprising, however, that answers change when people are asked an anchoring question prior to the

estimation. To put the problem more succinctly, we have the intuition that the 'objective' values which we assign to our mental representations *should* be the same in both contexts. Anchoring experiments might show that this is not the case. I want to push back on construing the spatial anchoring 'problem' in such a manner because it would also make judgments based upon visual information irrational. One "illusion" our visual system is susceptible to is that perception of color is influenced by other colors in the visual scene.[6] When viewed in isolation we may judge a splotch of color to be brown, but we judge the splotch as orange when it is surrounded by other color splotches. I do not know of any HB proponent who would claim that color illusions are problematic violations of the transitivity norm[7]. At best, color illusions show us that our visual system is not perfect. When we have no other way of identifying the color of an object, that is not external to our native mental systems, it is still perfectly reasonable to make judgments based upon the visual information. Similarly, when we represent objects in isolation we may judge them to be *x* inches long, but place it in a context with other objects and we judge it as *y* inches long. Since there is no other mental process which we can use to 'measure' the object, HB proponents should not claim that spatial anchoring effects are problematic violations either.[8]

Even if spatial anchoring does not violate either of the above norms, there is still a sense that spatial anchoring might be problematic. We might be worried that our spatial representation system is not a reliable source of information because it is not likely to be successful in generating the outcomes we want. I have argued earlier in this paper that a sticky spatial grid, as compared to a completely static grid, is an incredibly useful process which supports a wide range of (successful) behavior. Generally speaking, then, the spatial grid is akin to our visual system. Both systems generally produce information which is

---

[6] For examples see http://serendip.brynmawr.edu/bb/contrastcolor/

[7] Typically, when HB proponents mention color illusions it is meant to explain why heuristic processes are hard to de-bias. The de-biasing question is not a concern to discussions of rationality because the process is irrational (on HB's use of the term) regardless of whether telling people about experimental results gets them to make fewer errors. What is interesting about anchoring, to the HB proponent, is that we use the wrong *internal* process to generate our judgments in the first place.

[8] There is ample debate in philosophy about whether color exists as an actual property of an object or merely a psychological phenomenon (see (Byrne and Hilbert, 2003 for an overview)). If the analogy between our visual system and our spatial representation system holds at a deep level, then spatial anchoring may be a gateway to extending the color debate into perception of spatial properties as well.

accurate enough for most tasks and both systems are subject to certain inaccuracies. Moreover, the spatial representation system is highly adaptive to the wide range of environments we encounter and we would lose functionality if it was not sticky. Much like it is not irrational to rely upon our visual system as when making judgments which require handling visual information, it is not irrational to rely upon our spatial representation system when making judgments which require spatial information.

Perhaps the worry is not that our spatial representation system is problematic in the general sense but, rather, the use of the system is a problem when applied to a domain which it is not suited for. The claim might be, then, that our spatial representation system evolved from the need to operate in non-imagined activities. In *that* domain, it is very successful but when used in imaginative activities it should be considered unsuccessful. First, I find the claim that the use of imagination in order to deal with hypothetical situations, or solve spatial problems, constitutes a shift in domain highly implausible. It is implausible because it would suggest that imagining where someone is behind me (so that I could hand them my coffee mug while continuing to read the newspaper) is somehow very different from imagining how a different furniture arrangement would look in my living room.

Second, *even if* hypothetical or imagined situations *did* count as a separate domain, the evidence from anchoring experiments alone are insufficient to establish that the spatial representation system is unsuccessful in that domain. To claim that the spatial representation system is unsuccessful in the domain of imagined scenarios is not supported by a wealth of behavioral and empirical evidence. Behaviorally, we certainly employ our representational system during imagination with seemingly positive results. The fact that we can mentally rearrange our furniture or even tell if an elephant could comfortably fit in our living room suggests that we at least *think* the results are good enough to base behavior upon (see Gendler 2004 for a defense of extracting new information from imagination). The impression that imagined spatial scenarios are fairly accurate is also backed up by experimental work. For example, we are quite good at creating spatially coherent scenes from memory (Burgess, 2007) and our ability to estimate distances between objects in an environment where we have had no visual input is quite good (Fukusima, Loomis, and Da Silva, 1997). Additionally, there are many examples of successful

thought experiments in which we need to imagine the spatial relationships between objects in order for them to make sense (e.g., Galileo's falling object thought experiment or Feynman's Sticky Bead Argument). Therefore, there is ample reason to think that our spatial representation system should be considered successful even when we consider imagined scenarios as a domain which our spatial representation system has been conscripted to.

I have argued in this section for the idea that spatial anchoring effects need not (and should not) be considered the result of an irrational process. One of the reasons why anchoring research has mischaracterized spatial anchoring effects is because of the vague descriptions of the mental processes involved in anchoring. While I have no doubt that there are at least *some* instances of anchoring effects which stem from the use of an 'irrational' process, sorting those effects from effects generated by unproblematic processes is a project which needs to occur if the HB program to be successful at its own goals. Moreover, if the HB program is going to be useful in projects beyond the testing of rational theory, we need to have a greater clarity on which anchoring processes are truly worrisome.

8   Conclusion

Before concluding this paper, I wish to look at one area of further research. I have been explicit in limiting the Stick Space model to those anchoring phenomenon in a spatial domain. We might try and extend the theory to any process which uses the capacity for spatial representation, and that extension may allow us to make sense of other anchoring phenomenon. For example, there is some evidence that training on mental rotation tasks can significantly improve performance on numerical missing term problems (ex. 4+__= 11) (Cheng and Mix, 2014). This coincides with previous research that has shown a correlation between mental rotation ability and achievement in mathematics (Geary et al., 2000; Reuhkala, 2001). If it turns out that spatial representation plays a key role in our use of numbers (in the right sort of ways) then it is plausible that 'stickiness' in the representational space could produce cross-modal anchoring effects as well as numerical priming. I assume that our ability to perform mental rotation tasks with precision, is going to require being able to track several different points of the rotated object during rotation. That would suggest that we might use a spatial grid in the task. If a robust theory emerges for why it is that mental rotation is linked to mathematical concepts and ability, then it seems reasonable to think that we could turn around and apply that information to numerical anchoring effects. Moreover, the connection between our spatial representation system and numeracy would provide a path for the Sticky Space model to explain the cross-modal anchoring effects seen in Oppenheimer et al.

The extension is highly speculative and depends, not only, on the link between mental rotation and numeric concepts, but also on empirical and conceptual validation of the Sticky Space model. One way we might empirically test the Sticky Space model is based on the idea that subjects who are more adept at mental rotation and transformation tasks should be less susceptible to anchoring and less extreme in the response values. This would be the case since improved ability to scale and manipulate objects in the representational space should be indicative of fewer distortions in the scale of

that space. We might even have a testable prediction based along those lines right now. If anchoring is a result of distortions during scaling, then response times involving rotation of the target object should increase. They should increase because it will take more time to correct for the distortions induced in anchoring.

What I have attempted to do in this paper is to show that the spatial grid of our representational space that we use to generate information about the world can be the cause of certain anchoring effects. The range of anchoring phenomena that the model covers is intentionally smaller than the full range of phenomena that has been documented. Most researchers caveat their claims by saying multiple processes may be involved in any instance of anchoring. That seems correct, since not all study participants may be interested in expending the effort necessary to generate the best possible answer they can. In such cases, people may just pull the first number that comes to mind. That process, however, seems significantly different in comparison to those participants who actively consider how to best answer the question.

The problem I have tried to make for the major theories is that they gesture at some phenomenological point (e.g. 'feels right', 'seems reasonable', or 'semantically similar') in explaining anchoring, but that gesture is insufficient. Gesturing at the phenomenology is insufficient because it does provide us with the lower level processes that give rise to that phenomenology. Underspecification poses an additional problem when conducting research. Because the theories are underspecified in key areas, there is little ability to control for competing anchoring processes. If we want to design experiments to test the Scale Distortion model, we have to find a way to ensure that any observed effects are not attributable to SAM or anchor and adjustment models. If we do not have a clear idea of what 'semantic similarity' means, or how to control for the mechanisms that produce it, then we risk interpreting our results incorrectly. One benefit of a Sticky Space model is that the focus on low level abilities, like tracking the location and relative distance of objects in space, to begin differentiating

anchoring effects. In this case, it separates *spatial anchoring* effect from anchoring effects which might be produced in other ways.

The ability to differentiate anchoring phenomenon is not only useful for designing and interpreting experiments, but it also allows us to get a better grasp on when anchoring should, or should not, be considered a threat to rational decision making. Without proper differentiation, it is easy to mischaracterize anchoring effects as violations of the norms of rationality. It may be true that some forms of anchoring are 'irrational', in some sense of the word, but we cannot assume that *all* cases are. I have argued that spatial anchoring does not violate transitivity nor does it violate prohibitions on considering irrelevant information. Therefore, spatial anchoring should not been seen as being 'irrational' on the definition that proponents of the Heuristics and Biases program employ.

Moreover, unless we have information about rivers and animals that is stored in communicable forms, it is not clear that we have any other way to make the estimation task tractable. In other words, unless we already know the answer we are unable to generate estimates without having to resort to external means (tape measures, bathroom scales, Google, etc.). I have argued that our 'sticky' representation system is, for most intents and purposes, a reliable source of information from which to base judgments on. If it were not a generally reliable source, we would not be able to successfully perform a wide range of common behaviors. That assessment should not change merely because the mechanism has certain 'inaccuracies'. The analogy to our visual systems, in this case, is particularly apt. Our visual system is prone to certain 'illusions' which are similar to spatial anchoring, yet we do not think that relying upon visual stimulus is normatively problematic; especially when we do not have access to external tools to verify the information.

References

Bird, Chris M., James A. Bisby, and Neil Burgess. 2012. "The Hippocampus and Spatial Constraints on Mental Imagery." *Frontiers in Human Neuroscience* 6. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3354615/.

Brewer, Noel T., and Gretchen B. Chapman. 2002. "The Fragile Basic Anchoring Effect." *Journal of Behavioral Decision Making* 15 (1): 65–77. doi:10.1002/bdm.403.

Brun, Vegard Heimly, Trygve Solstad, Kirsten Brun Kjelstrup, Marianne Fyhn, Menno P. Witter, Edvard I. Moser, and May-Britt Moser. 2008. "Progressive Increase in Grid Scale from Dorsal to Ventral Medial Entorhinal Cortex." *Hippocampus* 18 (12): 1200–1212. doi:10.1002/hipo.20504.

Burgess, Neil. 2007. "Spatial Memory: How Egocentric and Allocentric Combine." Trends in Cognitive Sciences 10 (12): 551–557

Byrne, A., and D.R Hilbert. 2003. "Color Realism and Color Science." *Behavioral and Brain Sciences* 26: 1–44.

Byrne, Patrick A., David C. Cappadocia, and J. Douglas Crawford. 2010. "Interactions Between Gaze-centered and Allocentric Representations of Reach Target Location in the Presence of Spatial Updating." *Vision Research* 50 (24): 2661–2670.

Chapman, and Johnson. 1999. "Anchoring, Activation, and the Construction of Values." *Organizational Behavior and Human Decision Processes* 79 (2): 115–53. doi:10.1006/obhd.1999.2841.

Chen, Ying, Patrick Byrne, and J. Douglas Crawford. 2011. "Time Course of Allocentric Decay, Egocentric Decay, and Allocentric-to-egocentric Conversion in Memory-guided Reach." *Neuropsychologia* 49 (1): 49–60.

Cheng, Yi-Ling, and Kelly S. Mix. 2014. "Spatial Training Improves Children's Mathematics Ability." *Journal of Cognition and Development* 15 (1): 2–11. doi:10.1080/15248372.2012.725186.

Clearwater, Jonathon M., and David K. Bilkey. 2012. "Place, Space, and Taste: Combining Context and Spatial Information in a Hippocampal Navigation System." *Hippocampus* 22 (3): 442–54. doi:10.1002/hipo.20911.

Cosmides, L., and J. Tooby. 1996. "Are Humans Good Intuitive Statisticians After All? Rethinking Some Conclusions from the Literature on Judgment Under Uncertainty." *Cognition* 58: 1–73

Critcher, Clayton R., and Thomas Gilovich. 2008. "Incidental Environmental Anchors." *Journal of Behavioral Decision Making* 21 (3): 241–51. doi:10.1002/bdm.586.

Doeller, Christian F., Caswell Barry, and Neil Burgess. 2010. "Evidence for Grid Cells in a Human Memory Network." *Nature* 463 (7281): 657–61. doi:10.1038/nature08704.

Englich, Birte, and Thomas Mussweiler. 2001. "Sentencing Under Uncertainty: Anchoring Effects in the Courtroom." *Journal of Applied Social Psychology* 31 (7): 1525–51.

Epley, Nicholas, and Thomas Gilovich. 2001. "Putting Adjustment Back in the Anchoring and Adjustment Heuristic." *Psychological Science* 12 (5): 391–96.

———. 2005. "When Effortful Thinking Influences Judgmental Anchoring: Differential Effects of Forewarning and Incentives on Self-generated and Externally Provided Anchors." *Journal of Behavioral Decision Making* 18 (3): 199–212. doi:10.1002/bdm.495.

Frederick, Shane and Daniel Mochon. 2012. "A Scale Distortion Theory of Anchoring." *Journal of Experimental Psychology* 141 (1): 124-133. doi:10.1037/a0024006

Frederick, Shane, Daniel Kahneman, and Daniel Mochon. 2010. "Elaborating a Simpler Theory of Anchoring." *Journal of Consumer Psychology* 20 (1): 17–19. doi:10.1016/j.jcps.2009.12.004.

Fukusima, Sergio S., Jack M. Loomis, and José A. Da Silva. 1997. "Visual Perception of Egocentric Distance as Assessed by Triangulation." *Journal of Experimental Psychology: Human Perception and Performance* 23 (1): 86.

Galinsky, Adam D., Thomas Mussweiler, and Victoria Husted Medvec. 2002. "Disconnecting Outcomes and Evaluations: The Role of Negotiator Focus." *Journal of Personality and Social Psychology* 83 (5): 1131–40. doi:10.1037//0022-3514.83.5.1131.

Galletti, C., P.P Battaglini, and P. Fattori. 1995. "Eye Position Influence on The parieto-occipital Area PO (V6) of the Macaque Monkey." *European Journal of Neuroscience* 7: 2486–2501.

Geary, David C., Scott J. Saults, Fan Liu, and Mary K. Hoard. 2000. "Sex Differences in Spatial Cognition, Computational Fluency, and Arithmetical Reasoning." *Journal of Experimental Child Psychology* 77 (4): 337–53. doi:10.1006/jecp.2000.2594.

Gendler, Tamar Szabó. 2004. "Thought Experiments Rethought—and Reperceived." *Philosophy of Science* 71 (5): 1152–1163.

Gigerenzer, Gerd. 1996. "On Narrow Norms and Vague Heuristics: a Reply to Kahneman and Tversky." http://psycnet.apa.org/psycinfo/1996-01780-008.

Gigerenzer, Gerd, and U. Hoffrage. 1995. "How to Improve Bayesian Reasoning Without Instruction: Frequency Formats." *Psychological Review* 102: 684–704.

Green, Donald, Karen E. Jacowitz, Daniel Kahneman, and Daniel McFadden. 1998. "Referendum Contingent Valuation, Anchoring, and Willingness to Pay for Public Goods." *Resource and Energy Economics* 20 (2): 85–116.

Hafting, Torkel, Marianne Fyhn, Sturla Molden, May-Britt Moser, and Edvard I. Moser. 2005. "Microstructure of a Spatial Map in the Entorhinal Cortex." *Nature* 436 (7052): 801–6. doi:10.1038/nature03721.

Hassabis, Demis, and Eleanor A. Maguire. 2007. "Deconstructing Episodic Memory with Construction." *Trends in Cognitive Sciences* 11 (7): 299–306.

Jacowitz, Karen, and Daniel Kahneman. 1995. "Measures of Anchoring in Estimation Tasks." *Personality and Social Psychology Bulletin* 21 (11): 1161–66.

Just, Marcel Adam, Patricia A. Carpenter, Mandy Maguire, Vaibhav Diwadkar, and Stephanie McMains. 2001. "Mental Rotation of Objects Retrieved from Memory: a Functional MRI Study of Spatial Processing." *Journal of Experimental Psychology: General* 130 (3): 493

Kahneman, Daniel, and Amos Tversky. 1996. "On the Reality of Cognitive Illusions." http://psycnet.apa.org/journals/rev/103/3/582/.

Kosslyn, Stephen M., Thomas M. Ball, and Brian J. Reiser. 1978. "Visual Images Preserve Metric Spatial Information: Evidence from Studies of Image Scanning." *Journal of Experimental Psychology: Human Perception and Performance* 4 (1): 47.

Kosslyn, Stephen M., Giorgio Ganis, and William L. Thompson. 2001. "Neural Foundations of Imagery." *Nature Reviews Neuroscience* 2 (9): 635–42.

LeBoeuf, Robyn A., and Eldar Shafir. 2006. "The Long and Short of It: Physical Anchoring Effects." *Journal of Behavioral Decision Making* 19 (4): 393–406. doi:10.1002/bdm.535.

Mullally, Sinéad L., and Eleanor A. Maguire. 2013. "Exploring the Role of Space-defining Objects in Constructing and Maintaining Imagined Scenes." *Brain and Cognition* 82 (1): 100–107. doi:10.1016/j.bandc.2013.02.013.

Mussweiler, Thomas, and Birte Englich. 2005. "Subliminal Anchoring: Judgmental Consequences and Underlying Mechanisms." *Organizational Behavior and Human Decision Processes* 98 (2): 133–43. doi:10.1016/j.obhdp.2004.12.002.

Mussweiler, Thomas, and Fritz Strack. 2000. "The Use of Category and Exemplar Knowledge in the Solution of Anchoring Tasks." *Journal of Personality and Social Psychology* 78 (6): 1038–52. doi:10.1037//0022-3514.78.6.1038.

———. 2001. "The Semantics of Anchoring." *Organizational Behavior and Human Decision Processes* 86 (2): 234–55. doi:10.1006/obhd.2001.2954.

Northcraft, Gregory, and Margaret Neale. 1987. "Experts, Amateurs, and Real Estate: An Anchoring-and-Adjustment Perspective on Property Pricing Decisions." *Organizational Behavior and Human Decision Processes* 39: 84–97.

O'Keefe, John, and Neil Burgess. 1996. "Geometric Determinants of the Place Fields of Hippocampal Neurons." *Nature* 381 (May): 425–28.

Oppenheimer, Daniel M., Robyn A. LeBoeuf, and Noel T. Brewer. 2008. "Anchors Aweigh: A Demonstration of Cross-modality Anchoring and Magnitude Priming." *Cognition* 106 (1): 13–26. doi:10.1016/j.cognition.2006.12.008.

Pylyshyn, Zenon. 1973. "WHAT THE MIND'S EYE TELLS THE MIND'S BRAIN: A CRITIQUE OF MENTAL IMAGERY." *Psychological Bulletin* 80 (1): 1–24.

Quattrone, George A. 1982. "Overattribution and Unit Formation: When Behavior Engulfs the Person." *Journal of Personality and Social Psychology* 42 (4): 593.

Reuhkala, Minna. 2001. "Mathematical Skills in Ninth-graders: Relationship with Visuo-spatial Abilities and Working Memory." *Educational Pscyhology* 21 (4).

Samuels, Richard, Stephen Stich, and Michael Bishop. 2002. "Ending the Rationality Wars." *Common Sense, Reasoning, and Rationality*: 236–268.

Strack, Fritz, and Thomas Mussweiler. 1997. "Explaining the Enigmatic Anchoring Effect: Mechanisms of Selective Accessibility." *Journal of Personality and Social Psychology* 73 (3): 437.

Switzer, Fred, and Janeet Sniezek. 1991. "Judgment Processes in Motivation: Anchoring and Adjustment Effects on Judgment and Behavior." *Organizational Behavior and Human Decision Processes* 49: 208–29.

Todd, Peter M., and Gerd Gigerenzer. 2000. "Précis of Simple Heuristics That Make Us Smart." *Behavioral and Brain Sciences* 23 (05): 727–741.

Wallin, Annika. 2013. "A Peace Treaty for the Rationality Wars? External Validity and Its Relation to Normative and Descriptive Theories of Rationality." *Theory & Psychology*: 0959354313489369.

Wang, Ranxiao Frances, and Elizabeth S Spelke. 2002. "Human Spatial Representation: Insights from Animals." *Trends in Cognitive Sciences* 6 (9): 376–82. doi:10.1016/S1364-6613(02)01961-7.

Wegener, Duane T., Richard E. Petty, Kevin L. Blankenship, and Brian Detweiler-Bedell. 2010. "Elaboration and Numerical Anchoring: Breadth, Depth, and the Role of (non-)thoughtful Processes in Anchoring Theories." *Journal of Consumer Psychology* 20 (1): 28–32. doi:10.1016/j.jcps.2009.12.007.

Wilson, Timothy D., Christopher E. Houston, Kathryn M. Etling, and Nancy Brekke. 1996. "A New Look at Anchoring Effects: Basic Anchoring and Its Antecedents." *Journal of Experimental Psychology: General* 125 (4): 387.

Wong, Kin Fai Ellick, and Jessica Yuk Yee Kwong. 2000. "Is 7300 m Equal to 7.3 Km? Same Semantics but Different Anchoring Effects." *Organizational Behavior and Human Decision Processes* 82 (2): 314–33. doi:10.1006/obhd.2000.2900.

Zhang, Hui, and Arne Ekstrom. 2013. "Human Neural Systems Underlying Rigid and Flexible Forms of Allocentric Spatial Representation." *Human Brain Mapping* 34 (5): 1070–87.