Lost in Numbers?
Anchoring Effects in Advertising Claims and Product Information

Summary
According to anchoring theory, if unsure, human beings are predisposed to treat the first information they see as a starting point when making a judgement. This, often sub-conscious process, means random information can influence decisions in ways consumers are often unaware of. This paper tests this principle in advertising contexts to understand how anchoring may affect the way consumers interpret numbers within marketing messages. The results support the semantic priming and semantic anchoring models, which predict that random numbers will bias estimates when the wording of the 'anchor' is similar to the object of the estimate. We present evidence that this is the case even when the information is not directly relevant to the task. Contrastingly, no evidence is found to support the 'simple numeric priming' view of anchoring, which predicts that entirely abstract information can bias estimates.

1. Introduction

Anchoring is defined in the behavioural economics literature as the human predisposition to treat a provided number, or other value, as a starting point when making a judgement (Tversky and Kahneman, 1974). Research into anchoring disproportionately focuses on reference prices. An example would be the £450 original price with a line through it, which makes a designer handbag seem like a bargain at £100. This aspect of anchoring has been heavily researched and robustly demonstrated. On the other hand, price information is only one part of anchoring. All kinds of numerical values (Furnham and Boo, 2011) or even images (Madzharov and Block, 2010) can engender anchoring effects, which is why understanding non-price anchors is equally important (Yadav, 1994).

Anchoring bias can be extremely wide-ranging and even entirely irrelevant anchors in product information can influence consumer perceptions (Yun, 2010, Madzharov and Block, 2010). Such effects are well understood with regard to reference prices, but the rather less extensive research in other areas suggests anchors may have far broader implications in marketing and advertising contexts than existing Marketing research suggests.

Therefore, the purpose of this paper is to provide insight into how anchoring biases may affect the way consumers interpret numbers within marketing messages while perusing the retail aisle or browsing for products online. In particular, the study tests (a) whether a claim on a product label might affect consumers' estimates about adjacent products' characteristics; (b) whether arbitrary numbers in product information can influence guesses about the price of the product or about other product properties; and (c) whether implausibly extreme advertising claims are in fact more effective than bold but more plausible claims.

2. Literature Review

2.1. Origins of the Anchoring Effect

Anchoring is a bias exerted by initially presented values upon decisions made under uncertainty (Tversky and Kahneman, 1974). The theory challenges the assumptions of classical economics, which assumes that, given access to all the information, consumers make rational choices. In contrast, anchoring suggests that human beings have neither the time nor capacity to process all the information they are exposed to, and regularly treat the first

information they see as a starting point for attempting an estimate or making a judgement (ibid).

In the decades since the first anchoring effects were demonstrated (Tversky and Kahneman, 1974), numerous experiments have reported similar findings (e.g., Epley and Gilovich, 2001; McElroy and Dowd, 2007; Mussweiler and Strack, 2001). Experiments demonstrate that anchoring biases are particularly acute in cases where the true value is unclear (Van Exel et al., 2006).

Although some have questioned such experiments' external validity (Wu et al. 2008), anchoring research has been applied in a wide variety of settings from forecasting, management, negotiation and judicial sentencing to marketing and advertising (Furnham and Boo, 2011).

2.2. Anchoring Research in Marketing and Advertising

Anchoring research in marketing literature is mostly concerned with its influence on consumers' price sensitivity. Several studies show how reference prices are used effectively to influence consumers' estimates of the value of products and services (Biswas et al., 1993; Northcraft and Neale, 1987; Ariely et al., 2003). While price anchoring is well established, this is less the case for other potential applications (Yadav, 1994; Yun, 2010; Wilson et al., 1996).

Among limited research on non-price anchoring, Wansink et al. (1998) present empirical evidence that non-price anchoring in advertisements can exert a powerful, measurable influence on consumer buying behaviour. Furthermore, Madzharov and Block (2010) show that imagery on product packaging (such as the number of pretzels pictured on the bag) can anchor consumers' estimates of the amount of product contained inside − proving it is possible for biases to be recorded using anchoring cues on product packaging. Yun (2010) was able to bias participants' preferences simply by modifying the model number on a product. Yun reports that participants rated a fictitious 'Dell Studio 81' laptop more favourably than an identical laptop marked 'Dell Studio 18' laptop. These findings suggest that even entirely arbitrary anchors in product information may alter participants' evaluation of products.

Non-price studies generally follow a methodological framework distinct from studies concerned with price. This is important, because non-price related anchoring information is often secondary in comparison to price, which is a central preoccupation of the consumer. In many cases it is therefore preferable to introduce the anchor values incidentally instead of asking participants to focus on the anchor value (Critcher and Gilovich, 2008; Wilson et al., 1996). For example, while it may be appropriate to ask estate agents whether they would value a property at more or less than the reference price (given that prices are compared and discussed all the time in the real estate profession), it would be a different matter to ask a consumer if they would give a Dell laptop a higher or lower favourability rating than the product model number. The latter would lack external validity since it is far removed from the consumer's natural thought process when evaluating products. The fact that biases have been observed without prompting by the experimenter, suggests anchoring may be relevant in real world retail settings.

There is therefore considerable scope to build upon previous non-price studies. Specifically, more research is needed to understand the extent to which anchoring may distort consumers' evaluation of products. Few papers have researched this issue, which is surprising given the bold, numerical nature of some claims, for example:

- A deodorant that claims to last for up to 96 hours (Right Guard Xtreme Dry, 2015);
- An energy drink that claims to boost energy for up to five hours (5-Hour Energy drink range, 2015); and
- A moisturiser that claims to hydrate skin for up to 7 days (Garnier 7-Day body lotion, 2013).

While this has been explored outside the anchoring literature, studies have centred largely on rational decision making, for example establishing a link between brand reputation and consumer trust in evaluating bold or exaggerated claims (Goldberg and Hartwick, 1990) or the importance of evidence to support consumers to make informed choices (Xie and Boush, 2011; Marks and Kamins, 1988).

2.3. The Psychology of Anchoring

While anchoring effects have been repeatedly demonstrated over forty years, the underlying mechanisms are only partially understood.

Anchoring is in fact now thought to be 'an umbrella concept' for several mechanisms, which determine the effect depending on the context (Epley, 2004). Four key theories, which attempt to explain anchoring in different contexts include: (i) insufficient adjustment (Tversky and Kahneman, 1974); (ii) selective accessibility (Epley and Gilovich, 2001, 2005); (iii) attitude change (Wegener et al., 2001); and (iv) priming (Wong and Kwong, 2000; Mussweiler and Strack, 1999).

(i) Insufficient adjustment

Tversky and Kahneman (1974) initially proposed anchoring to be the result of a two-stage anchoring and adjustment process. Insufficient adjustment is primarily useful for understanding the anchoring effect in situations where anchors are self-generated (Epley and Gilovich, 2001). In marketing, an example would be a consumer in the market for a new car. He or she might estimate the price by adjusting upwards from the price of the previous model (Epley, 2004).

Experiments concerned with self-generated anchors normally measure the anchoring effect using a two-step question, known as the standard anchoring paradigm (Epley and Gilovich, 2006). Participants are first asked to estimate the anchor value. For example Wansink et al. (1998) presented shoppers with a special offer for a series of grocery items before asking them to self-generate an anchor in the form of the question 'How many units of this product do you usually buy?' This is followed by an absolute question. In this example it was 'How many units will you buy today?' The results are then compared against a control group (who are only shown the absolute question).

(ii) Selective accessibility

The insufficient adjustment thesis becomes difficult to sustain in some circumstances – for example, when the anchor is implausible (Kahneman, 2011). The 'selective accessibility' model redressed some of these gaps in understanding (Chapman and Johnson, 1999;

Mussweiler and Strack, 2001). Its premise is that, when attempting an estimate, our minds' first recourse is to recall any information available to support the anchor value presented – a process defined as hypothesis-consistent testing (Mussweiler and Strack, 2001). Thus, when participants must 'consider the opposite' (i.e. think of reasons why the anchor should be rejected) they present with reduced biases estimates (Chapman and Johnson, 1994). The selective accessibility model is particularly useful in explaining the effect of provided anchors.

Experiments exploring the effects of 'provided' anchors sometimes follow the standard anchoring paradigm formula in which the effect is measured by a two-part question. First, the participant is required to make a comparative assessment (e.g. 'Do you think the car is worth more or less than 40,000 German Marks?') Next, participants must make an absolute judgement such as 'How much is the car worth?' (Mussweiler et al., 2000).

Selective accessibility can also explain anchoring biases when the anchors are introduced incidentally. For example, consumers are not asked to consider the reference price in a store, although its presence influences perceptions; not only about a product's value, but also, by comparison, about the value of surrounding products (Nunes and Boatwright, 2004).

(iii) Attitude change

The 'attitude change' principle may help explain some aspects not addressed by other theories of anchoring. Findings by Wegener et al. (2001) suggest that extremely exaggerated or implausible anchors can be counterproductive i.e. less effective than a more modest anchor. They propose that this is because extremely implausible anchors trigger the mind to form counter-arguments that balance the influence of the anchor. For example, using the standard anchoring paradigm, participants were asked questions such as 'Was Ernest Hemingway older or younger than 2 years old when he wrote his first novel?' and found that the anchoring effect was often weaker than when the more plausible anchors were used (e.g. 16 years) (Wegener et al., 2001).

(iv) Priming

A fourth perspective is that anchors engender a priming effect. There is significant disagreement among scholars surrounding the nature of priming effects in anchoring. In particular, Strack and Mussweiler (1997) attribute anchoring with a 'semantic priming' effect, having found biases to be particularly strong when the anchor is on the same a scale as the target response. For example, a height value, known to be arbitrary (e.g. 150 metres), exerted a strong influence over estimates of the height of the Brandenburg Gate, but not over width estimates. Extending this principle to situations in which the object of the judgement is different, Frederick and Mochon (2012) show how estimates of a giraffe's weight were directionally biased after first estimating a racoon's weight.

However, research in other areas indicates that entirely abstract values can bias estimates. Some experiments have found that simply seeing an anchor prompts people to consider numbers close to the anchor value. In one study participants' estimates of how much they would spend at a restaurant were influenced by whether the restaurant was named ''Studio 17'' or ''Studio 97.'' Meanwhile, the number on an athlete's jersey biased percentage estimates about his performance in the next game. Seeing the random number is believed to

increase the accessibility of similar numbers, thereby influencing participants' estimates (Critcher and Gilovich, 2008). Findings by Chapman and Bornstein (1996) and Wong and Kwong (2000) support this view.

There appears to be some variety in the approach to experiments into priming effects, with some following the standard anchoring paradigm, more often associated with the selective accessibility model. However, it is open to question whether this two-step question format has sufficient external validity (Wu et al., 2008).

Indeed, studies into priming effects have used 'incidentally introduced' anchors (Yun, 2010; Critcher and Gilovich, 2008). In this way, when exploring priming effects, researchers have tended to introduce the anchor within an image or text without drawing attention to it or requesting a direct comparison. The anchoring bias is then measured by comparing responses between high and low anchor experimental groups. Since this method limits intervention by the researcher, it has been associated with less powerful effects than 'provided' anchors (Wilson et al., 1996).

2.4. Research Hypotheses

Informed with this literature review, a set of hypotheses were developed to test the extension of anchoring bias in advertising and product information contexts and how this may affect the consumer decision making process. Our first hypothesis will attempt to contribute to under-researched non-price anchoring research. In particular, given that a random social security card number has been proven to influence price estimates in an auction setting (Ariely et al., 2003), we will explore whether irrelevant anchors exert an effect in the consumer decision making process:

> *H1: When participants are shown a product and asked to estimate how much they would expect to have to pay for it, the price will be biased towards prominent but irrelevant numbers in the advertising.*

The second hypothesis will challenge the theoretical position that any number can cause anchoring bias. Research by Critcher and Gilovich (2008) succeeded in influencing estimates of an athlete's performance simply by altering the number on their jersey. In much the same way, we will test whether an entirely irrelevant number on a product label can influence the consumer decision making process about unrelated product characteristics:

> *H2: When participants are shown a product and asked to estimate one of its properties, estimates will be biased towards prominent but irrelevant numbers on the packaging.*

The third hypothesis will test whether a numeric advertising claim on one product will bias consumers' estimates of a characteristic of the non-related adjacent product when the message is semantically similar. Given that Yun (2010) found the model number on a laptop can influence perceptions of the product; we will test whether such affects can also occur between adjacent products. In our questions, the anchors will use similar wording as the target of the estimate, to attempt to achieve findings similar to studies in which semantically related anchors have resulted in powerful effects (Strack and Mussweiler, 1997; Mussweiler and Strack, 2001; Frederick and Mochon, 2012) by extending the research into the consumer decision making process area:

H3: When participants are shown a product and asked to estimate one of its properties, estimates will be biased towards prominent but semantically relevant numbers on non-related adjacent products.

The fourth hypothesis will also test whether a semantically relevant anchor can bias consumers' estimates of a characteristic of an adjacent product; however in this case the two products will be in the same product category and the anchor characteristic will be the same as the target characteristic. If demonstrated, this may have ramifications for advertisers, who would not want competitors to benefit from their products' claims. Similarly, brand managers may wish to consider anchoring biases when negotiating shelf position with retailers. Since the two products are similar in this case, it is anticipated that any observed bias will follow the 'hypothesis consistent testing' process described within the selective accessibility model (Mussweiler and Strack, 1999). Unlike semantic priming models of anchoring, this process is most applicable when consumers consciously consider the anchor due to its contextual relevance:

H4: When participants are shown a product and asked to estimate one of its properties, estimates will be biased towards prominent relevant numbers on similar adjacent products.

The fifth hypothesis will explore if anchoring impacts on the consumer decision making process when consumers encounter implausibly extreme advertising claims. Building on the 'attitude change' model by Wegener et al. (2001), it will be tested whether consumers are less affected when an extreme anchor is used than a bold but more plausible anchor. . Contrastingly, if an effect is observed it may lend credibility to selective accessibility proponents, who have demonstrated anchoring effects with implausibly extreme anchors (Chapman and Johnson, 1994; Mussweiler and Strack, 1999), but this time extending the research into the consumer decision making process:

H5: When participants are shown a product and asked to estimate one of its properties, estimates will be biased towards relevant, but improbably extreme claims on the packaging.

3. Research Methodology

Taking a quantitative approach, an online questionnaire was administered between April and May 2014, which is similar to previous anchoring studies (e.g. Sleeth-Keppler, 2013; Simmons and Leboeuf, 2010). CINT UK Limited, an accredited market research agency, distributed the questionnaires. CINT assigned its members of over 80 target group panels representing a balance of the UK consumer segments to complete the survey.

Although a non-probability sampling method was undertaken, quotas for gender and age helped to ensure heterogeneity. Furthermore, invitations were sent out proportionally according to the UK census, to ensure geographical balance in the sample. The broad range of consumer segments recruited led to a considerable diversity in our sample. Consequently, 156 UK adult consumers drawn from England, Northern Ireland, Wales and Scotland participated in the research (Table 1).

Table 1: Sample distribution

Gender
Male 52%; Female 48%
Age
18-24: 13%
25-34: 18%
35-44: 18%
45-54: 17%
55-64: 15%
65+: 19%
Occupation
Homemaker: 9%
Studies: 9%
Full time work: 26%
Part time work: 13%
Own business: 5%
Retired: 25%
Unemployed: 13%

Data collection was conducted in two phases. First, as recommended by Jacowitz and Kahneman (1995), formal research began with a calibration survey (n=50). This was largely the same as the subsequent experimental questionnaires, except the numerical advertising claims (anchors) were omitted from product images. Then, the estimates collected during the calibration phase were used to set anchor values within the product images for two experimental questionnaires (high and low anchor) at pre-determined percentiles (Jacowitz and Kahneman, 1995).

4.1. Questionnaire Design

The questionnaire followed a single factor, independent group design. The dependent variable was participants' estimates about a characteristic (e.g. price of a lamp) of a series of pictured products. The independent variable was a random numerical anchor in each image (e.g. a notice about the wattage of the lamp light bulb), manipulated across two conditions (high anchor and low anchor). Differences between the two groups were measured to test the hypotheses. Detailed information for scenarios used to test each hypothesis is presented in Appendix 1.

For Hypotheses 1 to 4, anchors were set at the 90^{th} percentile of the calibration group estimates in the high anchor questionnaire and at the 10^{th} percentile in the low questionnaire. Anchors were introduced incidentally (Critcher and Gilovich, 2008), meaning participants were not directed to consider the claims − as in the real world (Wu et al., 2008).

Hypothesis 5 tested an exaggerated claim in comparison to a bold, but more plausible claim. Therefore, the low anchor was set at the 90^{th} percentile (as per the high condition in H1-4). There is no fixed method for setting implausible anchors (Wegener et al., 2001). In our experiment the extreme high anchor was set at 3 times the 90^{th} percentile anchor value. The

21-hour claim used was dramatically higher than boldest energy drink claim found advertised in the UK (5-Hour Energy, 2014).

Hypothesis 5 was tested following the standard anchoring paradigm, in which participants directly consider the claim before entering an estimate (Tversky and Kahneman, 1974).

The research respondents were asked to rate the believability of the claim (1='Completely believable; 5= Completely unbelievable') before estimating the true average value. This ensured they understood they were required to critically appraise the claim when formulating an estimate rather than simply enter the claim value as an estimate.

The experimental body of the questionnaire, relating to the hypotheses under investigation, included five primary questions. For each primary question, participants had to estimate a characteristic of a pictured product, for example: 'How many hours would you expect the mouthwash to keep your mouth feeling fresh?'

As with a study by Epley and Gilovich (2001), participants were also asked to explain in their own words how they arrived at their guesses after entering each estimate (a 140 character limit ensured manageable data analysis). This approach helped to understand whether participants had considered the random numbers, without drawing attention to the anchors.

Products and brand names were fictitious, designed using graphics software. This was to avoid the effect of other factors, since, for example, brand trust has been shown to be a moderator of product claim effectiveness (Goldberg and Hartwick, 1990).

In addition to experimental questions, a set of moderating factor questions were designed. Past research suggests individual factors such as mood (Bodenhausen et al., 2000) and intelligence (Bergman et al., 2010) may affect biases so an individual factor of specific relevance to this study was investigated − advertising scepticism. Participants completed the nine-question, 5-point SKEP scale, a measure of this variable. The SKEP scale has performed well against several internal validity tests (Obermiller and Spangenberg, 1998). Responses were coded as 1=strongly agree (least sceptical); 5=strongly disagree. Further, previous studies suggest a link between relevant experience and lower biases; so we asked the participants to record how frequently they used similar products after entering each estimate on a 7-point usage frequency Likert item (1=Never, 7=Multiple times each day), (Horn et al., 2011). This data was not requested for H4 since the scale was not appropriate for the product.

4. Research Findings

4.1. Simple Numeric Priming Effects on Price Estimates (H1)

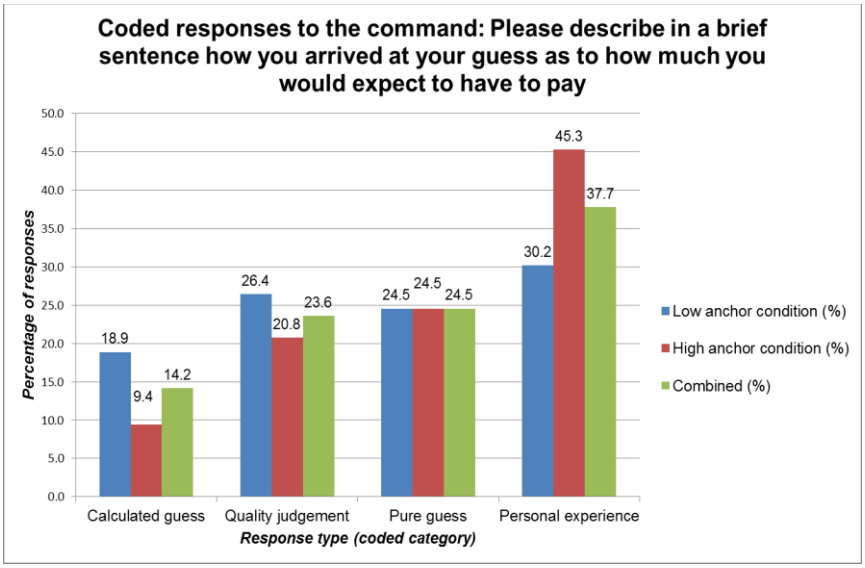
We tested a null hypothesis of 'no difference' against our hypothesis that when shown a product and asked to estimate how much they would expect to have to pay for it, participants' estimates would be biased towards irrelevant numbers in the advertising. There was no significant difference between the high and low anchor conditions (Mann-Whitney U=1244; two tailed p=.31; *n=53 in each group*). Therefore, the null hypothesis, that the two groups were the same, is retained.

Table 2: Summary statistics, Hypothesis 1

		Anchors		Medians		Transformed Medians	
	Calibration median	Low anchor	High anchor	Low anchor	High anchor	Low anchor	High anchor
Question							
Table lamp: How much would you expect to pay	£30	10 watt bulb	60 watt bulb	£20	£25	.196	.372

This outcome is consistent with respondents' explanations for their estimates. The analysis of their comments revealed four common response types as in Figure 1 (n=53 in each group). The largest proportion of participants reflected on their personal knowledge of the cost of similar products, or on past purchases. Around a quarter were unsure how they arrived at their estimate, which was coded as 'pure guess'. 'Quality judgements' were direct appraisals of the product without further elaboration. Examples include 'looks cheap' or 'nothing overly special'. The absence of any mention of the random anchor (the wattage of the light bulb) is in line with the interpretation that the irrelevant anchor was ignored.

Figure 1: Explanations for participants' estimates ('expect to pay', table lamp)



4.2. Simple Numeric Priming and Irrelevant Product Information (H2)

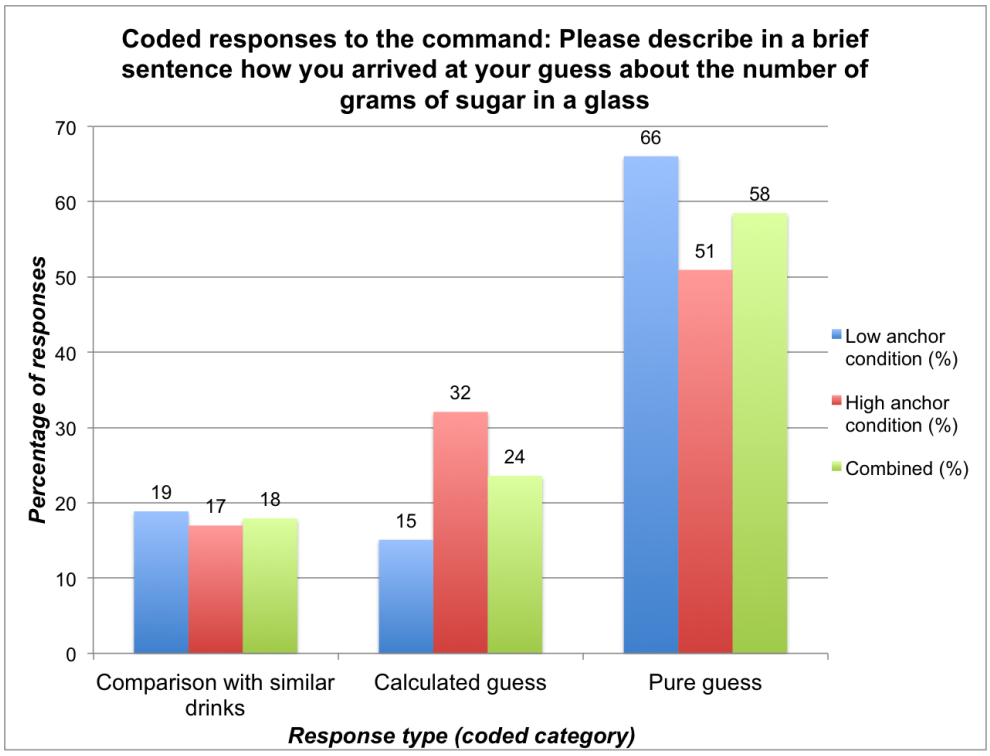
We tested the null hypothesis of 'no difference' against our hypothesis that when shown a product and asked to estimate one of its properties, participants' estimates would be biased towards prominent but irrelevant numbers on the packaging. There was no significant difference between the high and low anchor conditions (Mann-Whitney U=1293; two tailed p=.48; n=53 in each group). Therefore, the null hypothesis, that the two groups were the same, is retained.

Table 3. Summary statistics, Hypothesis 2

		Anchors		Medians		Transformed medians	
	Calibration median	Low anchor	High anchor	Low anchor	High anchor	Low anchor	High anchor
Question							
Lemon drink: The amount of sugar in an average glass	10 gr	2 prizes	86 prizes	8 gr	8 gr	.431	.431

The insignificant result is reaffirmed by the respondents' justifications for their estimates (n=53 in each group). The analysis of their comments revealed three categories (Figure 2). The majority of participants appeared to be totally unsure how much sugar went into a glass of lemonade, which was coded as 'pure guess'. 'Calculated guesses' made some attempt at rational estimation e.g. 'I'm assuming 6g of sugar per 100g, and a 200 ml glass'. 'Comparisons with similar drinks' were common, for example 'Coke has about 40 so I guessed it would be a little less than that'. However, none of the respondents mentioned the prize claim, indicating they did not consider the irrelevant number when making their estimates.

Figure 2: Explanations for participants' estimates (grams of sugar, Lemon Drink)



4.3. Semantic Priming Effects between Products in Different Categories (H3)

We tested our null hypothesis of 'no difference' against our hypothesis that when shown a product and asked to estimate one of its properties, participants' estimates would be biased towards prominent semantically related numbers on non-related adjacent products. Our analysis suggests that responses in the high condition were indeed considerably above those in the low anchor condition, and the difference was highly significant (Mann-Whitney U=990; two tailed p=.006; n=53 in each group).

Table 4: Summary statistics, Hypothesis 3

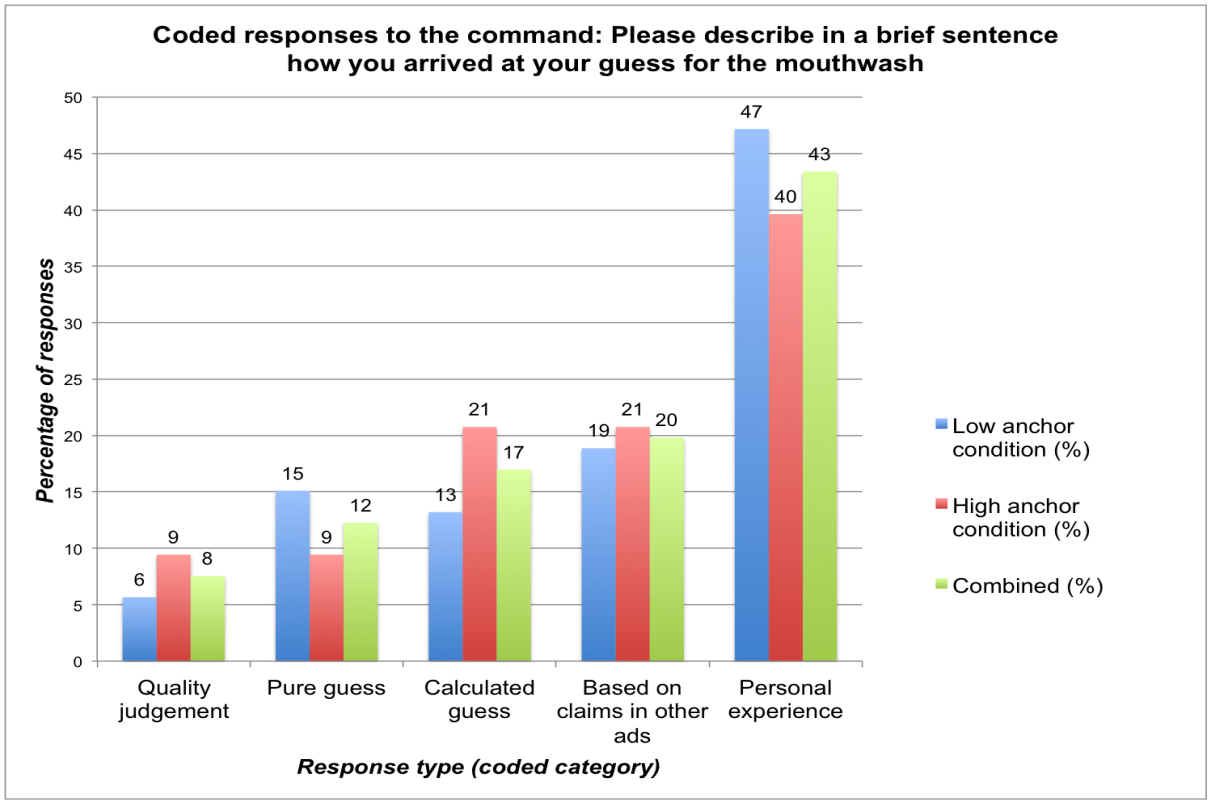
		Anchors		**Medians**		**Transformed medians**	
	Calibration median	Low anchor	High anchor	Low anchor	High anchor	Low anchor	High anchor
Question							
Mouthwash: Number of hours of freshness after use	4 hours	Up to 1 hour of freshness	Up to 12 hours of freshness	5 hours	12 hours	.568	.843

The bias recorded appeared to be primarily among the high anchor group. In fact, while the low anchor group median estimates were well below those in the high group (Table 4), they were slightly above the calibration group median (5 and 4 hours respectively), suggesting that low anchor group participants were not drawn towards the 1-hour claim. It should be highlighted that the negative skew in the data may also have contributed to the imbalance. Specifically, the 10th percentile of calibration group (our 1-hour low anchor value) was far closer to the median calibration group response (4 hours) than the 90th percentile high anchor (12 hours). Furthermore, because the low anchor was at 1-hour (near zero), participants could not estimate below the anchor value. Contrastingly, several participants estimated above the anchor value in the high condition. There was thus far less scope for anchoring biases in the negative direction. Nevertheless the overall highly significant anchoring bias is particularly striking given the two products were different.

The result is relevant to the underlying mechanisms of anchoring. The difference between this result and that of Hypothesis 2 is especially noteworthy. While both questions involved unrelated anchors, the anchor used on the adjacent deodorant spray in this experiment involved an anchor that was semantically the same as the target question about the mouthwash (the deodorant promised ' Up to *x hours* of *freshness'*, while participants were asked to estimate 'how many *hours* the mouthwash would keep your mouth feeling *fresh*'). On the other hand, the question asked on the lemonade ('how many *grams* of sugar..?') was semantically different from the anchor ('*x* number of *prizes* to be won').

The degree of influence goes beyond what participants appeared to have been conscious of. When asked how they arrived at their estimates, participants' responses were grouped into five categories as in Figure 3 (n=53 in each group).

Figure 3: Explanations for participants' estimates (hours of freshness, mouthwash)



'Personal experience' featured most heavily in participants' explanations. A typical response in this category would include 'It's how long my usual mouthwash lasts'. 1 in 5 participants also based their estimates on 'claims seen in other advertisements'. An example of a typical response in this category would be 'That is roughly what most advertisers say they last'. Guesses were common and two types of guess were identified. 'Pure guesses' were those who simply responded, for example 'I don't know' or 'it was a guess'. 'Calculated guesses' involved some attempt at rational estimation, for example 'I'm guessing that the effect would last until the next meal'. A relatively small number claimed to base their estimate on the quality of the product pictured. 'Quality judgements' such as 'Looks cheap' or 'looks good' accounted for 8% of total estimates.

What is interesting about these responses is the absence of any acknowledgement of the influence of the anchor on the adjacent product. In fact, participants did not mention the adjacent product at all across the two conditions. This would suggest that respondents were overall unaware of the bias exerted on their estimates by the adjacent anchoring claim, presented on the can of deodorant.

4.4. Selective Accessibility Effects between Products in the Same Category (H4)

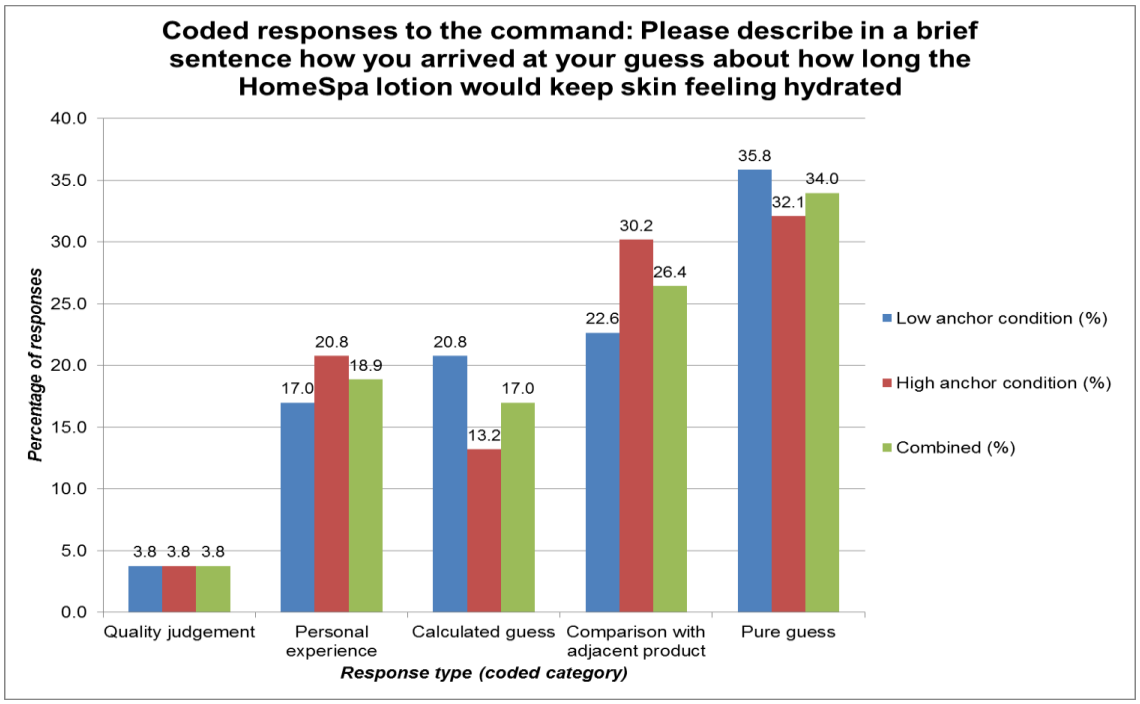
We tested our null hypothesis of 'no difference' against our hypothesis that when shown a product and asked to estimate one of its properties, participants' estimates would be biased towards prominent relevant numbers on adjacent products. We did so in the same manner as for H3. The only difference was that the two products were in the same category (body lotions) and the claim on the adjacent product was relevant. Responses in the high anchoring condition were considerably above those in the low anchor condition, and the difference between the groups was highly statistically significant (Mann-Whitney U=528; two tailed p=<.001; n=53 in each group). Therefore, we reject the null hypothesis.

Table 5: Summary statistics, Hypothesis 4

		Anchors		Medians		Transformed medians	
	Calibration median	Low anchor	High anchor	Low anchor	High anchor	Low anchor	High anchor
Question							
Body lotion: Number of hours hydration after use	5hr	1hr	24hr	2hr	12hr	.117	.725

Figure 4 shows the five types of justification for estimates identified (n=53 in each group). 'Pure guesses' accounted for 34% of responses across both groups. Examples of such answers include 'don't know' or 'I don't usually use those things so just a guess'. 'Comparison with adjacent product' was the second most diffuse response type reported. These answers were those that directly referenced the adjacent product. In some cases these respondents simply offered the same value as the adjacent product with a justification being 'Would be similar to Saints Treatment', while others appeared to take the anchor as a starting point and adjust away from it to arrive at their answer. One such response was 'it does not look as good quality as the one on the left so think would only last half as long'. 'Calculated guesses' demonstrated some attempt at deduction, for example 'some people have 2 showers or baths a day so would expect at least 12 hours'. Some 18.9% drew upon their 'personal experience' of using moisturisers, such as 'as soon as you get dressed and go out skin goes back to normal' or 'Because my own lasts a long time'. Finally a few respondents (3.8%) offered 'quality judgements' on the packaging by way of explanation, for example 'it doesn't make any special claims and it looks cheap'.

Figure 4: Explanations for participants' estimates (hours of hydration, body lotion)



4.5. Selective Accessibility vs. Attitude Change Theories In Relation To Exaggerated Anchors (H5)

We tested a null hypothesis of 'no difference' against our hypothesis that, when shown a product and asked to estimate one of its properties, participants' estimates would be biased towards relevant, but improbably extreme claims on the packaging. Results showed that while estimates in the high anchor condition of our experiment were higher than in the low condition, the difference fell just short of statistical significance (Mann-Whitney U=1107.5; two tailed p=.06; n=53 in each group). Therefore, the null hypothesis is retained.

Table 6: Summary statistics, Hypothesis 5

		Anchors		**Medians**	
	Calibration median	Low anchor	High anchor	Low anchor	High anchor
Question					
Energy drink: Hours of increased energy for the average person	2hr	7hr	21hr	2hr	2hr

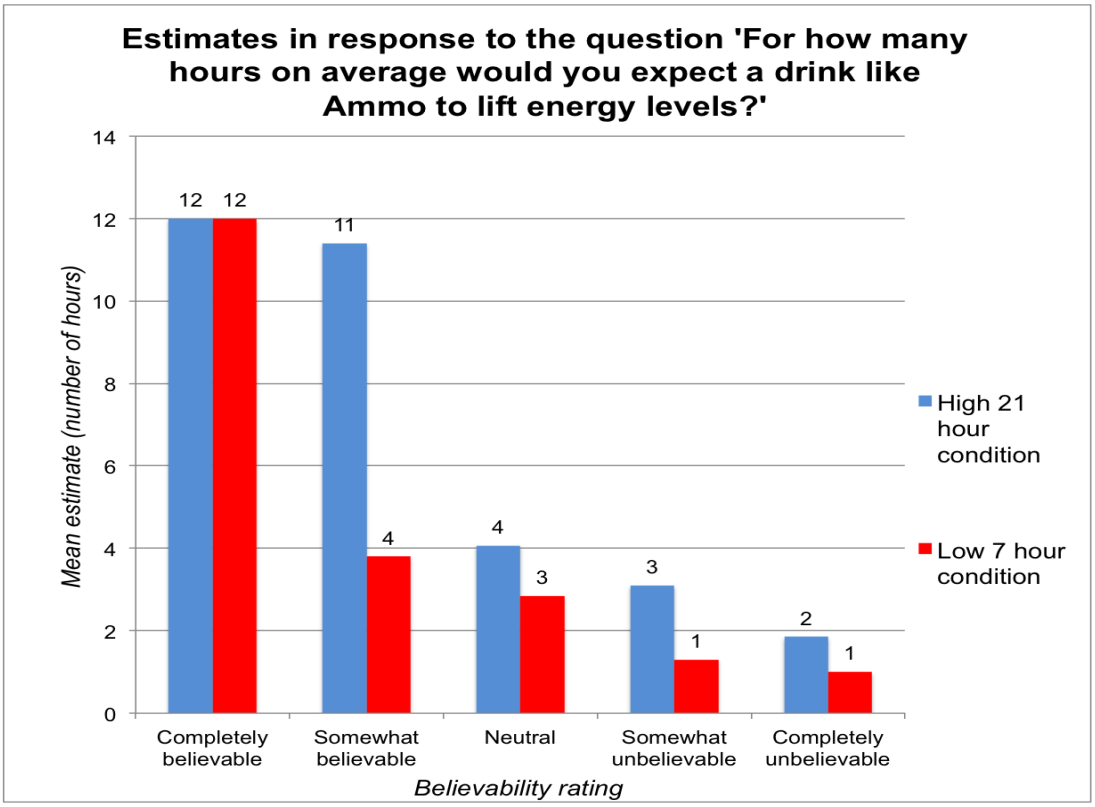
The characteristics of this 'threshold' result means further research is needed in order to position the finding squarely in the theoretical debates. A clear conclusion is difficult, not only because the finding is insufficiently significant to uphold the selective accessibility view, which has shown biases even for extreme anchors (Strack and Mussweiler, 1997), but also because the results do not show evidence for the alternative 'attitude change' model (Wegener et al., 2001). In our experiment the '21-hour' mean estimate was the highest (3.98 hours). Moreover, the mean in the 7 hour condition (2.5 hours) was in fact slightly lower than the calibration 'no anchor' group (2.8 hours). According to the attitude change model, we would anticipate the estimates in the more plausible anchor group to be higher than estimates in the extreme condition, which was not the case.

Content effects may partly explain the relatively high calibration group estimates. The group was not a pure control condition because, unlike the high and low groups, participants were not required to consider the believability of the claim before estimating. However, LeBoeuf and Shafir (2009) found that even forewarnings about anchors do not eliminate biases, suggesting this may not be a complete account.

Certainly, it does not explain why the bias between the high and low conditions fell short of significance. This outcome is surprising given the low anchor was more relevant to the question compared to H3 and H4 in which biases were found. A credible rationalisation is that there is simply little scope for anchoring biases for product claims when the range of plausible answers is very small (i.e. an energy drink is unlikely to lift energy for more than a couple of hours, while moisturisers may conceivably hydrate skin for 12 or even 24 hours). We have also observed that variance (and standard deviation) for responses to the energy drink question was lower than for all other questions across all groups, and therefore possibly resilient to manipulation. This would be in accordance with research connecting uncertainty with heightened bias (Van Exel et al., 2006). If this is the case, the experiment might produce significant effects using different products or claims.

However, a closer inspection of the data suggests that the extremely high anchor may have engendered some influence on estimates. Figure 5 shows that even though participants generally found the implausible 21-hour claim to be less believable than the 7-hour claim, their estimates tended to be higher regardless of believability ratings. In total 43% of participants found the 7-hour claim (low anchor) either somewhat or completely improbable, and this increased to 62% for the 21-hour (high anchor) claim. At the same time, the mean estimate was quite consistently above the high anchor condition regardless of whether participants judged the claim believable or not. The relatively consistent pattern presented in Figure 5 and the closeness of the result to significance, indicates further research is warranted into implausible product claims.

Figure 4: Believability ratings and mean estimates for the Ammo Energy Drink advertising claim



Note: Only one person in each group selected 'completely believable', limiting the relevance of this average.

4.6. Moderating Factors

Possible relationships between anchoring biases and four variables were tested. These included advertising scepticism levels, product usage frequency, age and gender.

For questions where an anchoring bias was observed, we tested whether advertising scepticism was correlated with the degree of anchoring bias. After transforming the estimates for the questions in which anchoring effects were observed into percentiles of the calibration group estimates, a Spearman's Rank Correlation Coefficient was calculated to observe whether participants' scepticism scores would reflect this trend. In fact, while this pattern was observed in most cases, no statistically significant correlations were found. Across the two questions, the mean correlations for high anchors and advertising scepticism were r =-0.05, n=53, two tailed p=0.51. For low anchors they were r =0.15, n=53, two tailed p=0.31.

Applying the same methodology we tested whether usage frequency of the products was correlated with the degree of anchoring bias among our participants, in cases where anchoring biases were observed. Again, if anchoring biases are stronger among those with low scores (in our case low product usage frequency), the correlation between estimates (transformed into percentiles of the calibration group responses) and product usage should be negative in the high anchor group and positive in the low anchor groups. This pattern was observed for the body lotion, but the correlation was not significant. For the mouthwash the correlation between product usage and anchoring estimates was significant in the high anchor condition, but it was a positive correlation (in the opposite direction than had been anticipated). The low anchor condition finding for the mouthwash was in the expected direction but was not significant. Overall no significant correlation between product usage frequency and anchoring bias was found. Across the two questions, the mean correlations for high anchors and product usage frequency were r =0.10, n=53, two tailed p=0.16. For low anchors they were r=0.19, n=53, p=0.18.

Our sample was closely matched to the age profile of the UK population as reported by the Office for National Statistics (2011). The Spearman's Correlation Coefficient was calculated between age (year of birth, provided by the survey company with participants' authorisation) and participants' transformed estimates. If anchoring biases are stronger among older participants (i.e. those with lower birth year values), the correlation between estimates and age should be negative in the high anchor group and positive in the low anchor groups; or vice versa if younger participants are more vulnerable to biases. While the pattern of results is consistent with a correlation between age and anchoring bias, the finding is not significant. Across the two questions, mean correlations for high anchors and birth year were r =-0.09, n=53, two tailed p=0.51. For low anchors they were r=0.10, n=53, two tailed p=0.46.

The experimental groups comprised near equal numbers of men and women (52% m, 48% f). In the high anchor condition there were 28 males and 25 females. In the low anchor condition there were 27 males and 26 females (n=106). The results of a Mann-Whitney U test did not show any significant difference between gender groups' estimates for either question.

5. Conclusion

The results of the hypotheses are as set out below in Table 7. Significant results (p=<.05) are in bold.

Table 7: Overview of findings, Hypotheses 1-5

Hypothesis	*n* (per group)	Target of estimate	Mann-Whitney *U*	*P (two tailed)*
1	53	Expected price of a table lamp	1244	.31
2	53	Grams of sugar in a glass of lemonade	1293	.48
3	**53**	**Hours of freshness after use, mouthwash**	**990**	**.006**
4	**53**	**Hours of hydration after use, body lotion**	**528**	**<.001**
5	53	Energy lift, hours (energy drink)	1108	.06

This paper aimed to contribute to under-researched non-price anchoring research, and investigated how bias in advertising and product information contexts may affect the consumer decision making process. We were unable to find evidence of simple numeric priming effects on price estimates. Participants' estimates about how much they would expect to pay for a lamp were unaffected by a prominent numerical claim about the wattage of the light bulb. Moreover, participants' explanations did not mention the abstract number.

Similarly, our investigation into simple numeric priming and irrelevant product information found no evidence that entirely irrelevant numbers on products can engender biases. A soda bottle label advertising the number of prizes to be won did not influence estimates about sugar content. Furthermore, participants did not mention the irrelevant number in their explanations.

Contrastingly, semantic priming effects between products in different categories were very strong. Estimates about how long a mouthwash would keep participants' feeling fresh were significantly higher when the neighbouring deodorant promised 'up to 12 hours of freshness' compared to '1 hour'. Importantly, participants' explanations did not indicate they were aware of the adjacent claim, suggesting the effect was subconscious.

Semantic anchoring effects were very strong between adjacent products in the same category. The strongest biases were observed between two body lotions, using a semantically linked claim. When participants were asked to estimate how many hours a pictured body lotion would hydrate their skin, their answers moved significantly towards the number of 'hours of hydration' promised on an adjacent competing brand. In their explanations, over 25% indicated they had adjusted their estimates based on the neighbouring product's claim.

Furthermore, the effect of exaggerated anchors was tested. We did not observe a strong bias when the anchor was highly implausible. Estimates about how many hours a can of energy drink would lift energy were not significantly higher when the claim read 'Lifts energy for up to 21 hours' compared to when it read '7 hours', suggesting participants were not influenced by the claim overall. However, the result of this experiment came very close to statistical significance (p=.06), implying further research into exaggerated claims is warranted.

Biases were very resilient where found. There was no indication that those holding more sceptical attitudes towards advertising were less biased. Nor did usage frequency of similar products or participant age correlate with bias. No significant difference between male and female estimates was observed.

5.1. Implications

The findings are suggestive of important possible conditions governing the effectiveness of anchors in product advertising contexts. Regulators might feel cautiously reassured that simple numeric priming is unlikely to be a viable form of manipulative advertising, since abstract anchors did not influence willingness to pay or estimates about product characteristics.

The most promising opportunity lies in semantic anchoring and semantic priming. The results have shown that when attempting to make estimates about products, our participants appeared to treat similar (though not necessarily relevant) information on adjacent products as reference points to aid guessing. This is interesting from the marketer perspective since it demonstrates how product information and advertising claims are evaluated relatively rather

than at face value. Moreover, it raises the prospect that a bold claim may 'rub off' on adjacent products by anchoring consumers' expectations to it. Thereby an inferior product could be perceived more effective than it would otherwise.

While the findings do not challenge the imperative to present product claims, they may have implications for shelf positioning. For example, generic or own-label brands may in fact benefit from being positioned adjacent to leading brands if the anchoring effect renders it more likely that they are perceived a 'close second'.

Likewise it may be advantageous to differentiate wording from competing brands. Semantically distinct advertising claims may disrupt the mental shortcut associated with semantic anchoring biases (Strack and Mussweiler, 1997). Brands may even seek to adopt proprietary words in order to effectively copyright advertising claims. Alternatively, generic brands may seek to model their claim language on the market leader. Such considerations may be worthy of incorporation into brand identity building, provided the findings can be consistently repeated.

Another possible implication relates to the plausibility of claims. While some research suggests that even improbable anchoring values can bias estimates (Chapman and Johnson, 1994; Northcraft and Neale, 1987), we were unable to demonstrate this. This calls into question the advertising practice of asserting implausible claims, using carefully worded phrases such as 'up to' to legally mislead consumers (BBC News Online, 2011). However, the near significance of this finding requires further research, before a definitive conclusion can be reached concerning such claims.

What is of potential public concern is that biases appear quite resilient to individual factors that might be expected to limit them. Consumers might be surprised that advertising scepticism levels were not significantly correlated with lower bias levels. Neither was experience of using similar products. Consumer groups may have a role to play in educating shoppers about biases, and could encourage mitigating strategies. For example, it would be worth exploring whether the 'consider the opposite' method (Mussweiler et al., 2000), might reduce biases. Of course this is only relevant in cases where participants are aware of the anchor, which, as we have seen, is not always the case.

References

5-hour energy range. (2015). *5-hour energy product description.* Available: http://www.5hourenergy.com/facts/. Last accessed: 25 February 2015.

Ariely, D., Loewenstein, G. and Prelec, D. (2003). "Coherent arbitrariness": Stable demand curves without stable preferences. *The Quarterly Journal of Economics*, 118(1), pp. 73-106

BBC News Online (2011). *Ofcom wants to ban misleading broadband speed ads.* Available: http://www.bbc.co.uk/news/technology-12611315. Last accessed: 25 February 2015.

Bergman, O., Ellingsen, T., Johannesson, M. and Svensson, C. (2010). Anchoring and cognitive ability. *Economics Letters* 107(1), pp. 66–68.

Biswas, A., Wilson, E. J. and Licata, J. W. (1993). Reference pricing studies in marketing: A synthesis of research results. *Journal of Business Research,* 27(3), pp. 239-256.

Bodenhausen, G.V., Gabriel, S. and Lineberger, M. (2000). Sadness and susceptibility to judgmental bias: the case of anchoring. *Psychological Science,* 11, pp. 320–323.

Chapman, G. B. and Bornstein, B. H. (1996). The more you ask for, the more you get: Anchoring in personal injury verdicts. *Applied Cognitive Psychology*, 10(6), pp. 519–540.

Chapman, G.B. and Johnson, E.J. (1994). The limits of anchoring. *Journal of Behavioral Decision Making,* 7(4), pp. 223–242.

Chapman, G. B., and Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational behavior and human decision processes*, 79(2), pp. 115-153.

Critcher, C. R. and Gilovich, T. (2008). Incidental environmental anchors. *Journal of Behavioral Decision Making*, 21(3), pp. 241-251.

Epley, N. (2004), "A Tale of Tuned Decks? A tale of tuned decks? Anchoring as accessibility and anchoring as adjustment. *The Blackwell Handbook of Judgment and Decision Making,* pp. 240-256

Epley, N. and Gilovich, T. (2001). Putting adjustment back into the anchoring and adjustment heuristic: differential processing of self-generated and experimenter-provided anchors. *Psychological Science,* 12(5), pp. 391–396.

Epley, N. and Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: differential effects of forewarning and incentives on self-generated and externally provided anchors. *Journal of Behavioral Decision Making,* 18(3), pp. 199–212.

Epley, N. and Gilovich, T. (2006). The anchoring-and-adjustment heuristic. Why the adjustments are insufficient. *Psychological science*, 17(4), pp. 311-318.

Frederick, S. W. and Mochon, D. (2012). A scale distortion theory of anchoring. *Journal of Experimental Psychology*: General, 141(1), p.124.

Furnham, A. and Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, 40(1), pp. 35-42.

Garnier. (2013). *7 day body lotion product description.* Available: http://www.garnier.co.uk/body-care/beauty/garnier/intensive-7-days/nutri-intensive-milk-cocoa-butter-250ml. Last accessed: 25 February 2015.

Goldberg, M. E. and Hartwick, J. (1990). The effects of advertiser reputation and extremity of advertising claim on advertising effectiveness. *Journal of Consumer Research*, pp. 172-179.

Horn, L., Elliott D. and Johnson A. (2011). Personal Care Product Survey Final Results Report, Portland University.

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

Kahneman, D. (2011). *Thinking, fast and slow.* 1st ed. New York: Farrar, Straus and Giroux, p.127.

LeBoeuf, R.A., Shafir, E. (2009). Anchoring on the "Here" and "Now" in time and distance judgments. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 35, pp. 81–93.

Madzharov, A. V. and Block, L. G. (2010). Effects of product unit image on consumption of snack foods. *Journal of Consumer Psychology*, 20(4), pp. 398-409.

Marks, L. J. and Kamins, M. A. (1988). The use of product sampling and advertising: Effects of sequence of exposure and degree of advertising claim exaggeration on consumers' belief strength, belief confidence, and attitudes. *Journal of Marketing Research*, 25(3), pp. 266-281.

McElroy, T. and Dowd, K. (2007). Susceptibility to anchoring effects: how openness- to-experience influences responses to anchoring cues. *Judgment and Decision Making*, 2(1), pp. 48–53.

Mussweiler, T. and Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology*, 35(2), pp. 136-164.

Mussweiler, T., Strack, F. (2001). The semantics of anchoring. *Organizational Behavior and Human Decision Processes*, 86, pp. 234–255.

Mussweiler, T., Strack, F. and Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin*, *26*(9), pp. 1142-1150.

Northcraft, G. B. and Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes*, 39(1), pp. 84-97.

Nunes, J. C. and Boatwright, P. (2004). Incidental prices and their effect on willingness to pay. Journal of Marketing Research, 41(4), pp. 457-466.

Obermiller, C. and Spangenberg, E. R. (1998). Development of a Scale to Measure Consumer Skepticism Toward Advertising. *Journal of Consumer Psychology*, 7(2), pp. 159-186.

Office for National Statistics (2011). 2011 Census: Aggregate data [computer file]. UK Data Service Census Support. Downloaded from: http://infuse.mimas.ac.uk. Last accessed: 15 September 2014.

Right Guard Website. (2015). *Right Guard Xtreme Dry product description.* Available: http://rightguard.com/locker/xtreme-dry. Last accessed: 25 February 2015.

Simmons, J. P., Leboeuf, R. A. and Nelson, L. D. (2010). The effect of accuracy motivation on anchoring and adjustment: Do people adjust from provided anchors? *Journal of*

Personality and Social Psychology, 99(6), pp. 917-932.

Sleeth-Keppler, D. (2013). Taking the High (or Low) Road: A Quantifier Priming Perspective on Basic Anchoring Effects. *The Journal of Social Psychology*, 153(4), pp. 424-447.

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

Tversky, A. and Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185 (4157), pp. 1124-1131.

Van Exel, N., Brouwer, W., van den Berg, B. and Koopmanschap, M. (2006). With a little help from an anchor. Discussion and evidence of anchoring effects in contingent valuation. *Journal of Socio-Economics*, 35, pp. 836–853.

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

Wegener, D. T., Petty, R. E., Detweiler-Bedell, B. T. and Jarvis, W. B. G. (2001). Implications of attitude change theories for numerical anchoring: Anchor plausibility and the limits of anchor effectiveness. *Journal of Experimental Social Psychology*, 37(1), pp. 62-69.

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

Wong, K. F. E. and Kwong, J. Y. Y. (2000). Is 7300 m equal to 7.3 km? Same semantics but different anchoring effects. *Organizational Behavior and Human Decision Processes*, 82(2), pp. 314-333.

Wu, C. S., Cheng, F. F. and Lin, H. H. (2008). Exploring anchoring effect and the moderating role of repeated anchor in electronic commerce. *Behaviour and Information Technology*, 27(1), pp. 31-42.

Xie, G. X. and Boush, D. M. (2011). How susceptible are consumers to deceptive advertising claims? A retrospective look at the experimental research literature. *Marketing Review*, 11(3), pp. 293-314.

Yadav, M. S. (1994). How Buyers Evaluate Product Bundles: A Model of Anchoring and Adjustment. *Journal of Consumer Research*, 21(2), pp. 342-353

Yun, L. (2010). Purely incidental numeric priming on product judgments: the effects of using high versus low product model numbers. *AMA Winter Educators' Conference Proceedings*, 21, pp. 248-249.

Appendix 1

Hypothesis 1:
Respondents were presented with a fictitious catalogue cut-out of the 'Luminosity Table Lamp'. The only information given was the name of the product and a prominent irrelevant statement containing our anchor: '60 watt bulb included' in the high condition and '10 watt bulb included' in the low condition. They were asked how much they would expect to pay for it (in GBP) if they saw it in a department store.

Hypothesis 2:
Respondents were presented a bottle of fictitious 'Juicy Lemon and Lime Drink' and asked to estimate how many grams of sugar a small glass of the drink contains. Our anchor was an irrelevant promotion on the label which read '86 prizes to be won' in the high condition and '2 prizes to be won' in the low condition.

Hypothesis 3:
Respondents were presented a can of fictitious 'Soul Fresh' deodorant beside a bottle of 'Aquasonic' mouthwash. They were asked how many hours they would expect the mouthwash to keep their mouth feeling fresh after use. An anchor was added to the adjacent deodorant. This was 'Up to 12 hours of freshness' in the high anchor condition and 'Up to 1 hour of freshness' in the low condition.

Hypothesis 4:
A similar scenario as Hypothesis 3 was presented to respondents. The only difference was that the two products were in the same category (body lotions) and the claim on the adjacent product was relevant. In the high anchor condition the claim 'Up to 24 hours hydration' was included on a bottle of fictitious 'Saint's Treatments Total Body Moisturiser'. In the low anchor condition this was 'Up to 1 hour of hydration'. Respondents were asked to estimate how many hours the adjacent 'HomeSpa Intense moisture lotion' would hydrate the skin after one application.

Hypothesis 5:
Respondents were presented a can of fictitious 'Ammo' energy drink. On it was the claim 'Lifts energy for up to 7 hours' (low anchor) or 'up to 21 hours' (high anchor). Since this was a test of implausible claims compared to high but more plausible claims, the low anchor was set at the 90th percentile of the calibration group responses (the same as the high anchor group percentile in H1-H4) and three times higher than this value for the implausibly high anchor. In addition to testing for difference between the two experimental conditions, calibration group responses were used to gauge the relative effectiveness of the two claims compared to 'no anchor'.

In high and low conditions, respondents were explicitly asked to consider the anchor value. This was necessary to verify whether respondents found the anchor improbable, by declaring on a five-point scale (from 'Completely believable to Completely unbelievable') how they felt about the claim that the energy drink lifts energy for *up to* 7 or 21 hours respectively. This also ensured that respondents knew they were expected to critically appraise the anchor rather than simply enter the claim value as their estimate in the follow up question. Respondents were then asked to estimate how many hours they would expect a drink like 'Ammo' to lift energy levels *on average*.