



UNIVERSITY OF WARSAW
FACULTY OF ECONOMIC SCIENCES

WORKING PAPERS

No. 40/2020 (346)

HONEY, MUGS AND CARICATURES: ANCHORS ON PRICES OF CONSUMER GOODS ONLY HOLD HYPOTHETICALLY

MAGDALENA BRZOSOWICZ
MICHAŁ KRAWCZYK

WARSAW 2020



Honey, Mugs and Caricatures: anchors on prices of consumer goods only hold hypothetically

Magdalena Brzozowicz, Michał Krawczyk

Faculty of Economic Sciences, University of Warsaw

Corresponding authors: mbrzozowicz@wne.uw.edu.pl, mkrawczyk@wne.uw.edu.pl

Abstract: We elicit willingness to pay for different types of consumption goods, systematically manipulating irrelevant anchors (high vs. low) and incentives to provide true valuations (hypothetical vs. real). On top of a strong hypothetical bias, we find that anchors only make a substantial, significant difference in the case of hypothetical data, the first experiments to directly document such an interaction. This finding suggests that hypothetical market research methods may deliver lower quality data. Moreover, it contributes to the discussion examining the mechanism underlying the anchoring effect, suggesting it could partly be caused by insufficient conscious efforts to drift away from the anchor.

Keywords: anchoring effect; hypothetical bias; WTP; experiment

JEL codes: D91, C93, M31

Acknowledgements:

The authors gratefully acknowledge the support of the National Science Centre, Poland, grant UMO-2017/27/N/HS4/02116.

1. Introduction

Many of the research methods commonly used to elicit consumers' willingness to pay (WTP) involve hypothetical questions; in other words, the responses have no financial consequences for the participant. Contingent valuation and conjoint analysis represent two popular examples of such methods. It is well known that hypothetical WTP tends to exceed that of actual WTP, an effect dubbed *hypothetical bias* (HB) in behavioural literature. If the size of the bias was broadly constant across individuals and circumstances, simply calibrating the hypothetical responses would be an easy fix. This, however, is unlikely to be the case. More generally, it is an important, plausible, and rarely investigated proposition that hypothetical data tends to be of lower quality.

In this study, we seek to verify one testable manifestation of this conjecture; namely, we examine whether the inclination to follow useless "hints" (otherwise known as the *anchoring effect*) is attenuated when the reported WTP has real-life consequences for the participant. Additionally, if this was the case, providing a low anchor could prove to be a simple, hitherto rarely if ever used, way to reduce or eliminate HB in marketing studies.

Of course, our study also contributes to the rich literature on the anchoring effect. Given that it represents a disturbing feature of human judgment and decision-making, it is certainly worth investigating whether it may be lessened when reported values have real consequences. Such an interaction would suggest that the effect may chiefly be understood in terms of insufficient incentives to modify the most directly available value, not in terms of humans' inherent inability to do just that. The practical implications thereof are substantial. Arguably, affecting (raising) the WTP is *the* goal of marketing and several ways to achieve this are clearly related to the idea of anchoring. Setting an initially high price and then selling at a discount is one common example; calling a large coffee *grande*, et cetera, to *disassociate* from the (low) price anchors of competitors is another.

Considering how the interaction between the anchoring effect and the HB affects WTP is thus important for the two areas of literature for both theoretical and practical reasons. Despite this, as we will explore, it has not yet been directly addressed in the empirical literature.

2. Literature review

HB can be defined as a systematic difference between valuation amounts observed in declarative research compared to actual WTP (Murphy and Stevens 2004). Previous meta-analyses have clearly shown that HB is ubiquitous and substantial. Summarising 77 studies, Foster and Burrows (2017) identified that the median hypothetical WTP values exceed those observed in real trials by as much as 39%. Of course, it would be an oversimplification to consider the value reported in incentivised trials as “real”. However, some studies have at least shown that they are better predictors of actual market behaviour (Chang et al. 2009).

As can be expected, the HB is to some extent moderated by the specific choice of research method. The literature shows substantial heterogeneity in this respect. Popular methods for obtaining the hypothetical value include direct open-ended questions (Doyon et al. 2015), choice experiments (Moser et al. 2014), and declarative Vickrey auctions (List 2001). Likewise, real values have been, among other ways, elicited using n -th price auctions (List 2003), the Becker-DeGroot-Marschak (BDM) procedure (Boyce et al. 1989), and the implementation of a randomly chosen decision card in the context of discrete choice modelling. Several investigations have refrained from eliciting specific WTPs altogether, instead only observing purchase/abstain decisions instead (Blumenschein et al. 1998). The general finding is that choice-based elicitation methods tend to reduce bias (Murphy et al. 2005; Penn and Hu 2018).

Several studies have also been conducted focusing specifically on the ways to de-bias the participants, for example, using cheap talk (Doyon et al. 2015; List 2001) and its relative, dubbed real talk (Alfnes et al. 2010). However, other studies have used various calibration techniques. Similarly, some researchers have focused on the effect of the sample (using students may contribute to the bias, Murphy et al. 2005) and the goods in question (finding HB to be weaker in private goods compared to public goods, List and Gallet 2001).

Another strand of literature which is relevant to the current research concerns anchoring (see Furnham and Boo 2011 for a review). This effect is obtained when a seemingly irrelevant number (an anchor) is shown to the participant and affects his or her numeric judgment—forecast, estimation, or evaluation. The effect was first explored by Tversky and Kahneman (1974); in their laboratory experiments, participants were asked whether the fraction of African countries belonging to the United Nations was higher or lower than the outcome on a wheel of fortune they observed¹, before providing their best guess. This final answer was strongly

correlated with the (obviously irrelevant, exogenously random) outcome on the wheel of fortune.

Tversky and Kahneman proposed that the participants anchor on the signal and adjust upwards if it is deemed too low and downwards if it is deemed too high. However, this adjustment tends to be insufficient, because it is stopped when the plausible range is reached. That is, a participant may have an idea that the fraction is probably somewhere between 40% and 70%. If an anchor of 30% is provided, the participant says that the target value is higher than the anchor and subsequently reports that it is approximately 40%. For an anchor of 80%, the participant indicates that the value is lower than that, perhaps somewhere close to 70%.

A prominent alternative explanation (although the two mechanisms are not mutually exclusive) was provided by Strack and Mussweiler (1997). These authors claim that anchors affect subsequent information retrieval: anchor-consistent information is more easily accessible, so that the ultimate response tends to be close to the anchor. The subsequent literature provided further theoretical insights that are beyond the scope of this paper. One important but contested proposition is that adjustment from the anchor requires conscious cognitive effort. Among its testable predictions, it has been reported that cognitive load amplifies the anchoring effect (Gilbert et al. 1988; Kruger 1999).

Diverse empirical studies have been orchestrated, with anchoring being observed for estimates of temperatures in Antarctica (Mussweiler and Strack 1999), the probability of a nuclear war (Plous 1989), judgment tasks (Englich and Mussweiler 2001), estimates of project duration (Lorko et al. 2018), lottery evaluation (Chapman and Johnson 1994), and price estimates (e.g., Mussweiler et al. 2000).

From the point of view of this study, the most important are the results of studies concerning declared preferences, especially WTP for specific goods. This strand of the literature is also vital from the economic and marketing viewpoints, as anchoring affects not only the average but also the whole distribution of WTP values and, consequently, the demand curve (Koçaş and Dogerlioglu-Demir 2013). Early examples include Johnson and Schkade's (1989) experiment involving certainty equivalents of abstract lotteries and studies by both Green et al. (1998) and Kahneman and Knetsch (1993) for WTP for public goods.

In an influential paper, Ariely et al. (2003) identified very strong anchoring effects on the valuation of unpleasant hedonic experiences (annoying sounds), as well as some common

consumer goods, such as small electronic items and wine. However, subsequent work suggested the limited robustness of these results. Bergman et al. (2010), Tufano (2010), and Maniadis et al. (2014) discovered weaker effects, whilst Simonson and Drolet (2004) only observed the moderate impact of anchors on WTP (but not Willingness to Accept, WTA) for consumer goods; Sugden et al. (2013) found that only plausible (not extreme) price anchors affect WTA (and, to a weaker extent, WTP); Fudenberg et al. (2012) reported minor effects for consumer goods but none for random lotteries; Alevy et al. (2015) found no effect on WTP for peanuts and collectible sports cards.

Then, while the anchoring effect is regarded to be robust in numeric judgment tasks (e.g., concerning general knowledge), its effect in valuation tasks was shown to be inconsistent. Yoon et al. (2019) examined the robustness of anchoring effects on preferences (measured by WTP). Based on an analysis of earlier studies and on their own experiments, they found that differences between experimental procedures cannot explain the inconsistencies in the existing literature, and that the anchoring effect is generally robust to experimental settings. Their findings were only partially confirmed in a recent meta-analysis of anchoring studies on WTA or WTP; Li et al. (2020), based on 53 studies from 24 articles, obtained an effect of moderate size. Moreover, they concluded that the strength and robustness of the phenomenon might not be as great as previously believed. Ioannidis et al. (2020), in the recent paper questioning the robustness of the anchoring effect on preferences, observed no priming effect on the valuation of a bottle of wine. They also performed a concise meta-analysis and demonstrated that the anchoring effect is weaker for familiar goods than for unfamiliar ones and also observed a null result of anchoring in studies with clearly uninformative anchors.

Some anchoring studies (e.g., Lorko et al. 2018) have used (non-trivial) real incentives, but none investigated how incentives to provide the “correct” value change the working of anchors by direct experimental manipulation, keeping other factors constant. That, of course, would be most relevant for our purpose. In the body of literature studies examining anchors and incentives, it is often claimed that incentives are not important, but in fact the evidence is scarce and often reported merely in passing. For example, Tversky and Kahneman (1974) devote exactly one sentence to the issue: “Payoff for accuracy did not reduce the anchoring effect”. We are thus not told the number of participants, which questions were incentivized, nor the nature of incentives, among other matters. Chapman and Johnson (2002) are equally enigmatic: “Chapman and Johnson ([still] unpublished data) used a procedure in which half the subjects

were paid according to the preferences implied by their judgments of simple lotteries. There was no reduction in anchoring when subjects' payoffs depended upon their responses". Likewise, Wilson et al. (1996) only sketched the design of Study 4, in which participants judged the number of surgeons and physicians in the phone book and were rewarded for accuracy.

The experimental procedures of Epley and Gilovich (2005) are reported more comprehensively. They investigated two types of anchors: standard ones, i.e., explicitly given by the experimenter, and self-generated ones (e.g., the year 1492 was a natural self-generated anchor for the question about the year in which the *second* European explorer landed in the West Indies provided by the experimenter). They argued (although Simmons et al. 2010 disagreed) that incentives (which were modest and manipulated between-subjects) could act differently depending on these two variations of the anchoring effect. Indeed, they observed that only self-generated anchors were weakened by monetary rewards.

It should be emphasised that all of the aforementioned studies on anchors and incentives involved judgment rather than valuation tasks. Moreover, these judgment tasks could typically be considered to be somewhat interesting or entertaining.

One plausible reason why (modest) real incentives might have made little difference was that responders were intrinsically motivated to guess correctly even in hypothetical cases. Indeed, one typically wants to answer correctly provided there is a well-defined "right" and "wrong". The WTP questions are different in this respect, because here, conversely, the response is largely a matter of personal preference. Incentives may thus make a difference in these kinds of questions, even if they do not in the case of standard judgment tasks.

In the aforementioned meta-analysis of studies measuring WTP and WTA, Li et al. (2020) reported that incentives do not attenuate the priming effect. However, a large majority of the included studies were hypothetical; the few that used incentives could also differ on other dimensions, reducing the statistical power to detect an effect.

The only paper we are aware of that provides a direct comparison of *valuations* of analogous objects under hypothetical and real scenarios is Jung et al. (2016). In a series of field experiments, the authors investigated the impact of anchors on the amounts submitted by customers under the pay-what-you-want (PWYW) system, some of which were replicated as hypothetical scenarios in the laboratory (studies 14a-d). They suggest that some moderators of anchoring may operate in the hypothetical scenario only and that hypothetical payments are

more sensitive to anchors than real ones. However, it must be emphasised that customers *cannot* be expected to offer their true WTP in a PWYW system. In this sense, while there were real incentives, there was no incentive compatibility, i.e., it was generally not in any participant's best interest to reveal their own preference by offering an amount corresponding to their actual WTP (Carson and Groves 2007). Specifically, offering little in a PWYW system could be a signal not only of not being interested in the product, but also of being selfish. It is perhaps not very surprising that when the participants get a hint that it is appropriate to offer a lot (a high anchor), they are willing to comply to avoid sending a signal that they are selfish, but more so when this can be done at no cost (as in the hypothetical condition).

3. Research hypotheses

We sought to verify the following main hypotheses:

H₁: HB influences the valuation of the good

Based on the extant literature, we expected that the participants we assigned to make a purely declarative decision would report a higher WTP for the good than those making actual purchasing decisions.

H₂: The anchoring effect influences the valuation of the good

We expected that participants who were shown a high anchor would report a higher WTP for the good than those shown a low anchor.

H₃: HB and the anchoring effect interact

We expected that anchoring would be stronger in the hypothetical case than in the case of actual purchasing decisions. In other words, we expected that a low anchor would reduce HB compared to a high anchor.

4. Design and procedures: the elements common to all the experiments

To test these hypotheses, three experiments were conducted. In each of them, the participants were asked to state their WTP for a presented product.

In each experiment, four different treatments were used in a 2x2 design: RealLow (real transactions, low anchor), RealHi (real transactions, high anchor), HypoLow (hypothetical valuation, low anchor), and HypoHi (hypothetical valuation, high anchor).

Under hypothetical conditions, the technique of directly eliciting the WTP value was used (a simple request: “give the maximum price at which you would be willing to buy the presented product”), with the participants understanding that their valuation was declarative only. By contrast, the incentive-compatible BDM technique (Becker et al. 1964) was used in the Real conditions. The BDM procedure is operationally efficient and commonly used in laboratory experiments, but also enables experimenters to elicit consumers’ WTP in typical purchase settings in the field (Wertenbroch and Skiera 2002). In this procedure, we asked participants to name the maximum price they would be willing to pay for the presented product, before drawing the transaction price from a pre-specified distribution (in the form of small cards placed in a box). Participants were not informed about the lower or upper bounds of the range from which the exogenous price was drawn, because these could represent alternative anchors (Bohm et al. 1997). If the participant’s offer was higher than or equal to the drawn price, the participant was required to buy the product at the drawn price. If the offer was lower than the drawn price, the transaction was not executed. The weakly dominant strategy in this procedure is to state the true WTP (Wertenbroch and Skiera 2002).

5. Experiment 1

Experiment 1 was a paper-and-pencil laboratory experiment in which participants were asked to state their WTP for a voucher for a black-and-white caricature or portrait in A4 format, painted by an experienced artist. They could acquire this voucher for themselves or for somebody else. It was valid for three months. We predicted that most participants would not have frequently purchased such goods. They would thus find it difficult to judge how much they would like the product or how much it would normally cost. Similarly, we expected this type of product would be sufficiently susceptible to anchoring.

The assignment to a Real vs. Hypo condition was randomised at the session level for the sake of convenience and to exclude the possibility of treatment contamination, whereas a Hi or Low anchor was randomly assigned to each individual within a session. At the beginning of the session, participants were informed about the basic principles of the experiment (see

Appendix A for instructions) both orally and in writing. The voucher was then presented to them and some examples of the artist's works were shown. Participants were asked whether they would pay 20 PLN (Low anchor) or whether they would pay 80 PLN (High anchor) for the voucher. In the Real condition, the participants then took part in a BDM procedure. The BDM procedure was implemented for three randomly chosen participants in each session (about 10-20% of participants), who paid for the vouchers with their own money. At the end of the experiment, the participants were asked to fill in the post-experimental questionnaire covering, among other things, their sociodemographic characteristics and interest in art; see Appendix B for the translation and Appendix C for variable labels. In the Hypothetical condition, the scheme was analogous, but the BDM procedure was not used; the participants simply provided their valuations in response to an open-ended hypothetical question.

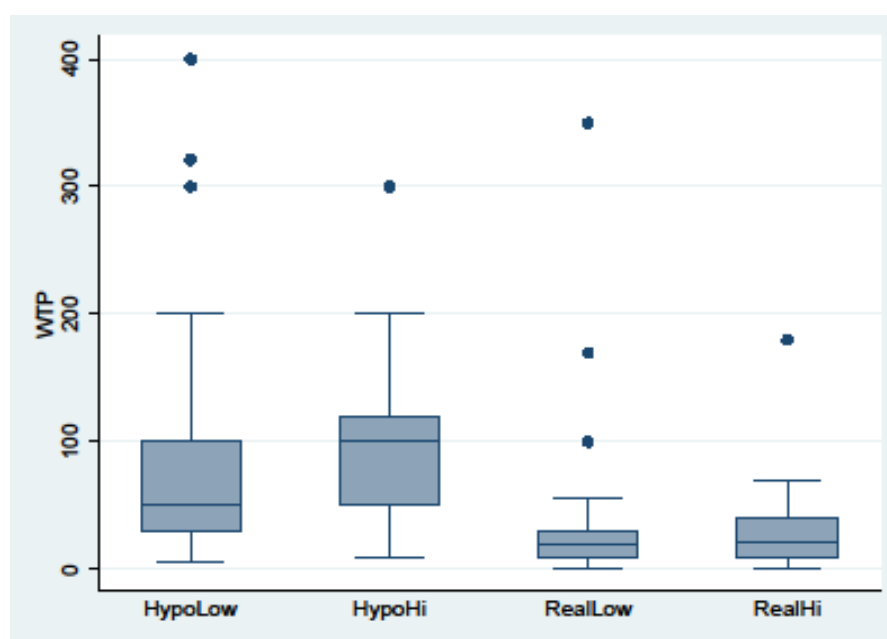
The experiment was conducted at the Faculty of Economics at a large and nationally prestigious university in Central and Eastern Europe. Because participants were recruited immediately after their classes and the sessions took approximately 15 minutes, no show-up fee was deemed necessary. In total, 218 local students took part, 53% of which were female. The mean age was 19.5 years; a typical participant was a single person with no employment, in a good financial situation.

5.1. Results of Experiment 1

The mean value of WTP for the voucher in the whole sample was equal to 58.32 PLN (12.88 EUR), with a median value of 40 PLN and a standard deviation of 65.83 PLN (for histogram, see Figure 4 in Appendix D). These aggregate measures mask substantial differences between the treatments, however. Because outliers, which are typically found in (hypothetical) WTP data, can severely distort our analysis, we first plotted the leverage and squared normalised residuals, see Figure 5 (Appendix D). It was clear that five observations should be dropped. One could debate whether observation no. 8 should be dropped as well; however, it made no difference to the outcomes. Unsurprisingly, dropping these five (or six) observations had no bearing on the Mann-Whitney tests, which will subsequently be presented. We started our investigation by comparing WTP values in each treatment, see Table 1 and Figure 1.

Table 1 Experiment 1: WTPs by Treatment (in PLN)

	HypoLow	HypoHi	RealLow	RealHi
Mean	82.08	98.10	30.04	29.03
Median	50	100	20	21
Standard dev.	86.83	56.94	50.46	28
N	51	52	58	56

Fig. 1 Experiment 1: WTPs by Treatment (in PLN)

Notes: The blue rectangle represents the middle 50% of the data (from the first quartile to the third), the line inside the box shows the median (the second quartile). The whiskers represent the top and bottom 25% values, excluding outliers, which are represented by dots.

From a visual inspection alone, the mean WTP was definitely lower in the Real condition than in the Hypothetical condition. The anchor made a difference in the Hypothetical treatments, but not in the Real treatments. We verified these observations using non-parametric Mann-Whitney U tests. First, we compared treatments with a Hypothetical vs. Real valuation to test for HB (separately for the low and high anchors). The difference between the WTP values in the HypoLow vs. RealLow treatments was statistically significant ($z = 5.486$, $p < .0001$), and likewise for the HypoHi vs. RealHi treatments ($z = 7.075$, $p < .0001$). Thus, we observed a clear HB. In the next step, we compared treatments with low and high anchors (separately for

hypothetical and auction valuations) to verify the anchoring effect. The difference between the WTP values in the HypoLow and HypoHi treatments was statistically significant ($z = -2.737$, $p = 0.0062$). By contrast, there was no difference in the case of real valuations ($z = -1.145$, $p = 0.2520$). In light of these results, we can state that the anchor affected the hypothetical valuation of vouchers (raising the mean WTP value by approximately 20%) but had no impact on valuations in the case of real transactions.

To verify that this pattern was not due to random differences in the innate characteristics of the four experimental groups and to explicitly test for the interaction, we conducted a regression analysis (see Appendix H for the variable labels). (We estimated the Ordinary Least Squares regression models with the WTP value as the dependent variable. All of the models included the experimental conditions (*hypothetical* and *high_anchor*). In model (2), we also included the interaction between the two. In model (3), we additionally controlled for certain sociodemographic characteristics such as gender, age, or place of origin. Model (4) further included factors that the participants took into consideration during the valuation process, e.g., the size of the portrait or caricature and the technique used by the artist. Finally, in model (5), we also included participants' self-reported interest in art and their perceived market price of the caricatures and portraits, as well as the extent to which the participant liked the example work shown to them previously and whether (s)he intended to offer the voucher for the portrait/caricature as a gift, should (s)he buy one.

Table 2 Experiment 1: Regression Table: WTP Values

	(1)	(2)	(3)	(4)	(5)
Hypothetical	53.669***	41.949***	41.903***	21.762***	24.714***
high_anchor	15.412***	4.600	5.006	8.085	11.051*
hypo#high_a		23.162**	23.119**	23.341**	15.185
Male			11.275*	4.771	7.874
Age			-1.023	1.581	1.430
Relationship			13.031**	6.057	4.346
Unemployed			13.001*	8.573	9.451*
City			-2.721	0.574	0.982
fin_situation			2.644	4.379	0.487
technique_e				13.170**	12.469**
format_e				4.391	-0.382
time_e				2.574	2.082
examples_e				-11.816**	-5.841
artist_e				19.528***	16.473**

needs_e				-23.777***	-20.207***
finance_e				4.175	4.914
others_e				12.508	13.131
Examples					17.716***
Gift					1.917
Art					8.133
price_caricature					0.181**
price_portrait					0.081
Cons	19.069***	24.426***	25.245	-6.734	-42.244
N	212	212	212	210	204
R-sqr	0.3022	0.3150	0.3501	0.4850	0.5825
F	45.25	31.89	12.09	10.64	11.48
Prob>F	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

* p<.10, ** p<.05, *** p<.01

We confirmed that the interaction between our experimental dimensions, *hypo#high_a*, plays an important role: in three of the four specifications in which it was included, it was significant at the 5% level (while the main effect of the anchor was not). The only exception was model (5), in which, among others, the *price_caricature* and *price_portrait* variables were included, which may have mediated the effect. Indeed, when a high anchor was given, the participants could plausibly update and increase their beliefs about market prices, subsequently affecting their WTP. We confirmed this conjecture using the Preacher and Hayes (2004) bootstrapped test of mediation, based on the Sobel-Goodman mediation test.¹

The remaining significant variables in our models affected the WTP in natural ways (in particular, it was lower in participants who had doubts if they needed the presented product at all and higher in those who liked the artist's works on display).

Because the data is not normally distributed, we applied yet another test of the interaction between the anchoring effect and the HB, namely the permutation test for analysis

¹ We considered four potential mediators: *art* (1-if the participant is interested in art, 0- in all other cases), *finance_e* (1- if the participant took into consideration in the valuation process how much (s)he can afford to pay, 0 – otherwise), *price_caricature* and *price_portrait* (the perceived market price of the caricature (portrait). In the case of the *price_caricature* variable, both the direct effect of the anchor and the indirect effect of the mediator were statistically significant in the hypothetical condition (for indirect effect $z=1.78$, $p=0.038$ and for direct effect $z=1.80$, $p=0.035$ in one-tailed tests, appropriate for our directional hypotheses). In the real condition, neither effect was statistically significant. For the *art*, *finance_e* and *price_portrait* variables in the hypothetical treatments, the indirect effect was not statistically significant (but the direct effect of the anchor was); in real treatments, neither the direct nor indirect effects were statistically significant.

of variance (permutation ANOVA). Again, the interaction between the two variables of interest was statistically significant; Table 3 illustrates the results.

Table 3 Experiment 1: Permutation ANOVA

Source	SS	df	F	parametric P>F	permutation p> F
hypothetical	151125	1	83.394	<0.0001	0.0002
high anchor	13809	1	7.62	0.0063	0.0060
hypothetical#high_anchor	7074	1	3.9	0.0495	0.0460
Residuals	376932	208			

To summarise, a large HB was found. Further, the presence of the incentives robustly interacted with the anchor. The anchoring effect was observed in the hypothetical answers, but not when real money was at stake.

The sample size of Experiment 1 was relatively small and comprised only students. It also featured just one type of good. To investigate the robustness of our results, we conducted two further studies.

6. Experiment 2

This study was a field experiment conducted in one of the largest shopping centres in Central Europe, Westfield Arkadia in Warsaw in September 2020. The participants were asked to state their WTP for a ceramic mug painted by an experienced artist. This product is potentially useful and desirable for customers of all ages and genders. Moreover, such mugs widely vary in price depending on the painting technique used, design, author's artistic talent, etc. We could thus expect it to be susceptible to anchoring, as participants would not know the actual market price of the valued good.

Before the experiment, we carried out an online survey in which the participants ($N=52$, diversified in terms of age and sex) were asked to pick the three most attractive designs of the 11 options presented. For the main experiment, we selected the set of three mugs (see Figure 6 in Appendix E) that maximized the fraction of the pilot study participants that picked at least

one of the three. This fraction was as high as 80%. Moreover, all but one of these participants rated the mugs eventually used in the main experiment as a “3” or “4” on a 1-4 scale.

The selected mugs were presented at a professionally prepared stand, located in a part of the shopping center with a very high footfall. The stand operated during the opening hours of the shopping center. The customers passing by, encouraged by the display of presented products, the banner promising a chance to win a prize (a 100 PLN Westfield Arkadia gift card), and sometimes a queue, approached the stand and were randomly assigned to individual experimental conditions.

In each treatment, at the beginning the participants were orally informed about the basic principles of the experiment (see Appendix F for instructions). We then showed them the mugs, enclosing additional information about the product and the artist. The participants were asked to choose which of the three mugs that they liked the most. They were then asked whether they would pay 10 PLN (in the Low anchor condition) or whether they would pay 60 PLN (in the High anchor condition) for the selected mug. Then, the participants were asked to state their WTP. Under hypothetical conditions, the technique of directly eliciting the WTP value was used, the participants were informed that their valuation was declarative only. By contrast, the incentive-compatible BDM technique (Becker et al., 1964) was used in the Real condition. For our participants' convenience, we accepted both payment in cash and by card. At the end of the experiment, the participants in each treatment were asked to complete a short post-experimental questionnaire providing additional information on their socio-demographic characteristics (see Appendix G for the questionnaire) and performed a task involving recreating on a touch screen a shape they briefly saw. Sufficiently precise drawings were rewarded with a 100 PLN gift card.

A typical experiment took approximately 10 minutes. In total, 786 shopping center customers (aged 16 and over) took part. The research sample was naturally diversified in terms of age, sex, education, etc. Roughly 53% of participants were female, with a mean age of 32 years.²

6.1. Results of Experiment 2

The mean WTP value for the ceramic mug across the sample was equal to 28.45 PLN (6.28 EUR), with a median value of 20 PLN and a standard deviation of 26.35 PLN (for histogram,

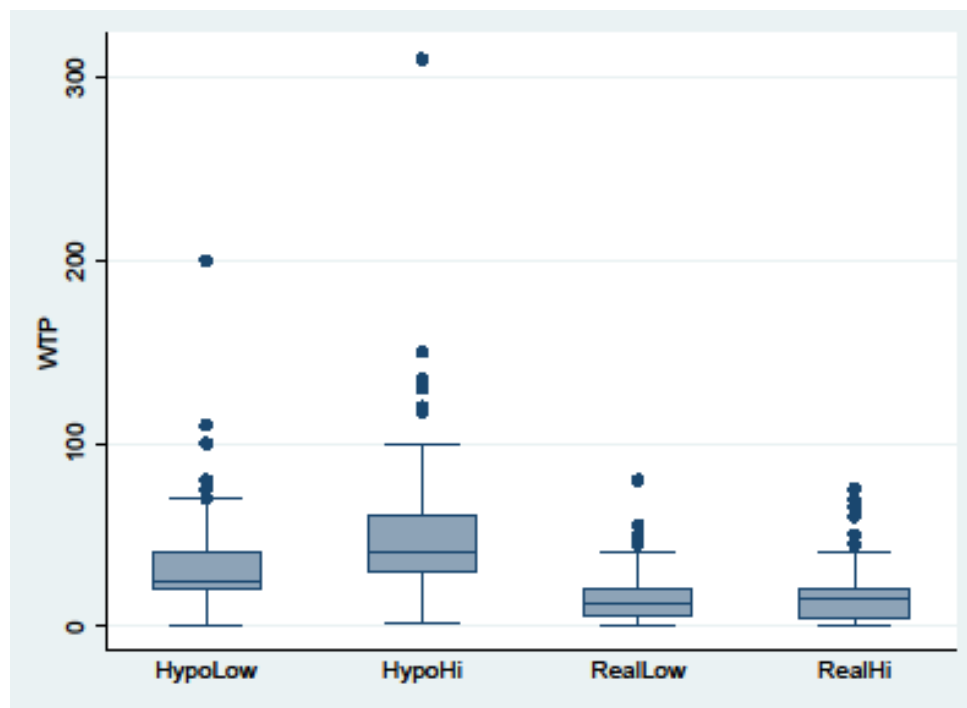
²¹This is close to the mean age of Westfield Arkadia visitors; before the pandemic, it was equal to 37.5 years but it has been lowered by the pandemic.

see Figure 7 in Appendix E). One outlier was removed using the same procedure as in Experiment 1 (observation no. 226 with a WTP value of 550 PLN). Again, we first compared the WTP values in each treatment; see Table 4 and Figure 2.

Table 4 Experiment 2: WTPs by Treatment (in PLN)

	HypoLow	HypoHi	RealLow	RealHi
Mean	32.38	49.39	15.48	16.19
Median	25	40	12	15
Std. Dev.	22.60	33.96	12.54	13.55
N	200	197	198	191

Fig.2 Experiment 2: WTPs by Treatment (in PLN)



Again, from a visual inspection alone, the mean WTP is higher in the Hypo condition than in the Real one. The anchor affects WTP values in treatments with hypothetical valuations, but not in treatments with real transactions. As in Experiment 1, we verified these observations using non-parametric Mann-Whitney U tests. Again, we compared treatments with a Hypothetical vs. Real valuation to test for HB. The differences between the WTP values in the HypoLow vs. RealLow treatments were statistically significant ($z = 9.897, p < .0001$), and likewise in the HypoHi vs. RealHi treatments ($z = 13.219, p < .0001$). Thus, we can

state that HB affects the valuation of the presented mugs. In the next step, we compared treatments with low and high anchors to verify the anchoring effect. The differences between the WTP values in the HypoLow and HypoHi treatments were statistically significant ($z = -6.615, p = < 0.0001$). By contrast, there was no difference in the case of real valuations ($z = -0.284, p = 0.7765$). The anchor affected the hypothetical valuation of vouchers (raising the average WTP values by about 53%) but had no impact on valuations in the case of real transactions.

We once more verified these findings using OLS regression models; see Table 5. Models 1-3 were analogous to the specifications in Experiment 1. Model (4) additionally included the extent to which the participants liked the presented mugs, as well as whether participants intended to offer the mug as a gift. Finally, in model (5), we also included the experimenter who carried out the particular experiment and which mug was selected by participants.

Table 5 Experiment 2: Regression Table: WTP Values

	(1)	(2)	(3)	(4)	(5)
Hypothetical	24.958***	16.903***	15.734***	14.940***	14.833***
high_anchor	8.925***	0.702	0.826	1.084	1.141
hypo#high_a		16.298***	16.698***	16.627***	16.704***
Male			-3.180**	-2.852*	-2.693*
Age			-0.276***	-0.286***	-0.289***
Education			-14.078***	-13.561***	-13.516***
Unemployed			-4.978***	-4.957***	-4.943***
Gift				0.790	0.848
very_nice				5.947***	6.149***
Experimenter					-0.067
mug_elephant					-2.122
mug_cat					-1.628
Cons	11.446***	15.483***	40.136***	36.794***	38.124***
N	785	785	779	779	779
R-sqr	0.2543	0.2782	0.3451	0.3577	0.3587
F	133.30	100.33	58.04	47.58	35.71
Prob>F	0.0000	0.0000	0.0000	0.0000	0.0000

* p<.10, ** p<.05, *** p<.01

In all the models allowing for the interaction (*hypo#high_a*), it was highly significant, while the main anchoring effect was not.

We also identified that participants with secondary or higher education reported higher WTP values than participants with primary or vocational education; older participants reported lower WTP values than younger ones and employed participants reported higher WTPs than unemployed ones. Unsurprisingly, WTP was also higher in participants who very much liked the presented mugs.

All the presented results were robust to the exclusion of inconsistent responses, i.e., those in which the reported WTP was lower than the anchor price that was accepted or higher than the anchor price that was rejected. This is not surprising given that only 2.29% of all responses showed this characteristic. While inconsistencies were more common in the Hypothetical condition than in the Real one, the difference was not statistically significant.

In the last step, as in the previous experiment, we applied the permutation ANOVA. Again, the interaction between the anchoring effect and HB was statistically significant; see Table 6 for the results.

Table 6. Experiment 2: Permutation ANOVA

	SS	df	F	Parametric P(>F)	Permutation P(>F)
hypothetical	28358.19	1	56.353	0.0000	0.0002
high_anchor	47.89	1	0.0951	0.7578	0.7650
hypothetical*high_anchor	13029.48	1	25.8923	0.0000	0.0002
Residuals	393013.83	781			

7. Experiment 3

Experiment 3 was also conducted in Westfield Arkadia shopping center in September 2020, but it did not overlap with Experiment 2. As the experiments were anonymous, we cannot verify how many people participated in both but given that there were few cases of perfectly matching demographic characteristics, we conclude that the overlap of the samples was very small indeed.

The design of Experiment 3 was very similar to that of Experiment 2; the key difference was that, instead of hand-painted mugs, we offered jars of flavored honey. After analyzing the popularity of various types, we selected three types: raspberry, garlic & ginger, and cocoa (see

Figure 8 in Appendix I). These types of honey are rarely accessible in shops, so we expected that most customers would be unaware of the market price of the offered product.

We used the same stand and location as in Experiment 2. We implemented the same 2x2 design as before to elicit the WTP for the type of honey they were most interested in. Before the participants were asked to state their WTP, they were again questioned whether they would pay 20 PLN (Low anchor) or whether they would pay 80 PLN (High anchor) for the selected honey.

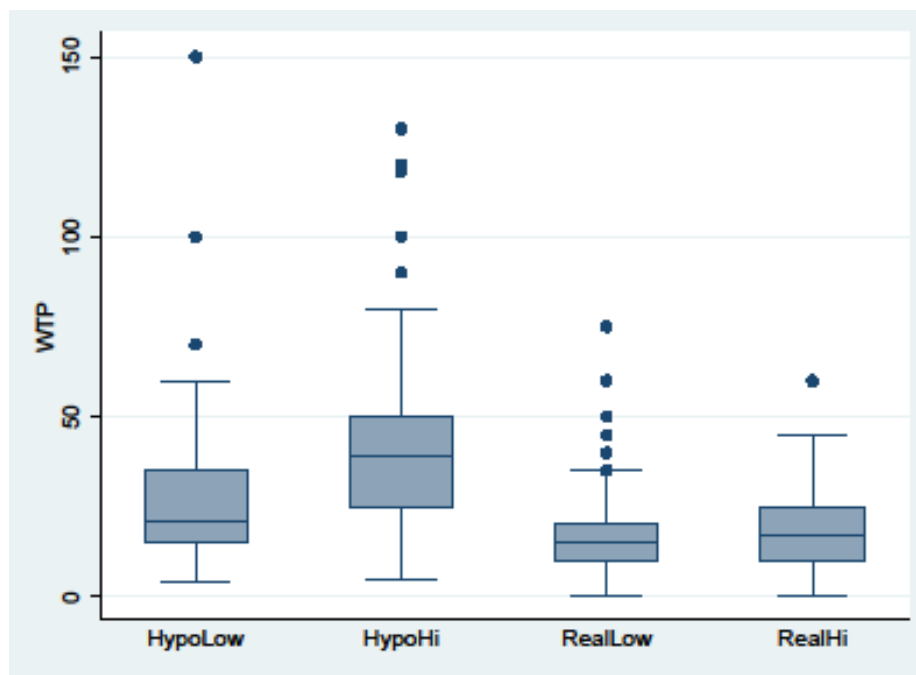
The experiment took approximately 10 minutes per participant. In total, 799 shopping center customers took part; 54% of participants were female and the mean age was 31 years.

7.1. Results of Experiment 3

The mean value of WTP for a jar of flavored honey across the sample was equal to 26.01 PLN (5.74 EUR), with a median value of 20 PLN and a standard deviation of 19.53 PLN (for a histogram, see Figure 9 in Appendix I). One outlier was removed using the same procedure as in Experiment 1 (observation no. 56 with WTP value of 2000 PLN). As in Experiments 1 and 2, we started our analysis by comparing the WTP values in each treatment; see Table 7 and Figure 3.

Table 7 Experiment 3: WTPs by Treatment (in PLN)

	HypoLow	HypoHi	RealLow	RealHi
Mean	28.00	41.41	15.76	17.96
Median	21	39	15	17
Std. Dev.	19.81	22.21	10.33	10.31
N	204	205	196	193

Fig. 3 Experiment 3: WTPs by Treatment (in PLN)

Similarly to before, WTP values in hypothetical treatments were higher than in real treatments and the anchor affected the valuation in the Hypo condition only. The differences between the WTP values in HypoLow vs. RealLow and in HypoHi vs. RealHi were highly significant ($z = 8.944, p < .0001$; $z = 12.625, p < .0001$), a clear HB. As for the anchoring effect, this time it was significant not only in the Hypo condition ($z = -7.856, p = < 0.0001$) but also in the Real condition ($z = -2.837, p = 0.0046$), although in the latter it was much weaker (48% vs. 14%).

As in previous experiments, to explicitly test for interactions between the effects under investigation, we estimated the OLS regression models with the same dependent variable as in Experiment 2 (WTP value). Models 1-3 were very similar to those of Experiment 2. In model (4), we additionally included the extent to which the participants liked the presented honey and how much they liked honey in general, as well as whether they intended to offer the honey as a gift. Model (5), again, included the dummy for the experimenter who carried out the particular experiment and which flavour was selected by participants. Table 8 presents the results.

Table 8 Experiment 3: Regression Table: WTP Values

	(1)	(2)	(3)	(4)	(5)
hypothetical	17.833***	12.242***	12.131***	11.663***	11.892***
high_anchor	7.941***	2.193	2.267	2.488	2.662
hypo#high_a		11.214***	11.206***	10.799***	10.525***
male			1.620	1.763	1.306
age			-0.118***	-0.116***	-0.116***
higher_edu			-1.229	-0.854	-0.717
unemployed			-3.749***	-3.658***	-3.550**
gift				3.919***	3.690***
attractive				6.348***	6.239***
likes_very_much				1.645	1.906
experimenter					1.345
honey_ginger					1.860
honey_cocoa					4.830***
cons	12.911***	15.762***	20.440***	13.710***	11.518***
N	798	798	791	791	789
R-sqr	0.2508	0.2714	0.2872	0.3130	0.3275
F	133.06	98.59	45.08	35.54	29.03
Prob>F	0.0000	0.0000	0.0000	0.0000	0.0000

The interaction between the two variables of interest was highly significant whenever it was included. The other significant variables in our models influenced the valuation in similar ways as in Experiment 2.

Again, the results were robust to the exclusion of inconsistencies. This time, they were slightly more numerous, representing 4.13% of all responses. The prevalence was once again higher in the hypothetical condition: (7.84% in HypoLow vs. 4.08% in RealLow and 3.90% in HypoHi vs 0.52% in RealHi); the latter difference was significant ($z = 2.267, p = .0234$).

The interaction effect was also confirmed in the permutation ANOVA, see Table 9.

Table 9. Experiment 3: Permutation ANOVA

	SS	df	F	Parametric P(>F)	Permutation P(>F)
hypothetical	14980.1	1	53.676	0.0000	0.0002
high_anchor	467.7	1	1.676	0.1958	0.1928
hypothetical*high_anchor	6268.1	1	22.460	0.0000	0.0002
Residuals	221592.3	794			

8. Conclusions

Although our three experiments involved two different settings and target populations and three different products, their results were remarkably consistent. We have strongly confirmed Hypothesis 1: HB influencing the valuation of a good. One reason for this tendency is that practical considerations may receive insufficient scrutiny when dealing with hypothetical choices. In our case, participants in the Hypo condition (of Experiment 1) were significantly less likely to state that they considered whether they actually needed the good than those in the Real condition (instead, they reported having devoted relatively more attention to the technique, the description of the artist, and examples of her works). Social desirability and experimenter demand effects may also have played some role. As a result, the participant may have overstated their actual WTP in the hypothetical condition. Hypothesis 2, predicting that anchoring influences the valuation of a good, was partially confirmed, as the effect was only found in the hypothetical condition (except for a weak effect in Experiment 3); consequently, Hypothesis 3 (the interaction) was strongly confirmed. This finding appears to be in opposition to the strand of research implying that the standard anchoring mechanism has little to do with insufficient mental effort (Epley and Gilovich, 2005). This theory would lead to the incorrect prediction of the weak effect of providing financial incentives to report truthfully. It also contradicts the aforementioned inadequately documented evidence, supposedly demonstrating that anchors continue to work when responses are real, which can be found in previous literature.

To the extent that susceptibility to anchoring is a bad sign concerning data quality, our findings speak out against the validity of hypothetical methods. This is corroborated by our additional observation of the greater prevalence of inconsistent responses in the hypothetical condition. Indeed, if anchors have much more impact on hypothetical than real decisions, providing a low one could be a promising tool in curbing HB, which could selectively reduce the reported valuations of those responders who tend to overstate them compared to their real WTP.

References

- Alevy, J. E., Landry, C. E., & List, J. A. (2015). Field Experiments on Anchoring of Economic Valuations. *Economic Inquiry*, 53(3), 1522–38.
- Alfnes, F., Yue, C., & Jensen, H. H. (2010). Cognitive Dissonance as a Means of Reducing Hypothetical Bias. *European Review of Agricultural Economics*, 37(2), 147–63.
- Ariely, D., Loewenstein, G., & Prelec, D. (2003). ‘Coherent Arbitrariness’: Stable Demand Curves without Stable Preferences. *Quarterly Journal of Economics*, 118(1), 73–106.
- Bergman, O., Ellingsen, T., Johannesson, M., & Svensson, C. (2010). Anchoring and Cognitive Ability. *Economics Letters*, 107(1), 66–68.
- Blumenschein, K., Johannesson, M., Blomquist, G. C., Liljas, B., & O’Conor, R. M. (1998). Experimental Results on Expressed Certainty and Hypothetical Bias in Contingent Valuation. *Southern Economic Journal*, 65(1), 169–77.
- Bohm, P., Lindén, J., & Sonnegård, J. (1997). Eliciting Reservation Prices: Becker–DEGroot–Marschak Mechanisms vs. Markets, *The Economic Journal*, 107(443), 1079–1089.
- Boyce, R. R., Brown, T. C., McClelland, G. H., Peterson, G. L., & Schulze, W. D. (1989). Experimental Evidence of Existence Values in Payment and Compensation Contexts. In *Proceedings of the Joint Meetings of the Western Committee on Benefits and Costs of Natural Resource Planning (W-133)*, 305–36.
- Brewer, N. T. & Chapman, G. B. (2002). The fragile basic anchoring effect. *Journal of Behavioral Decision Making*, 15(1), 65–77.
- Carson, R. T., & Groves, T. "Incentive and informational properties of preference questions." *Environmental and resource economics*, 37(1), 2007, 181–210.
- Chang, J. B., Lusk, J. L., & Norwood F. B. (2009). How closely do hypothetical surveys and laboratory experiments predict field behavior? *American Journal of Agricultural Economics*, 91(2), 518–34.
- Chapman, G. B., & Johnson, E. J. (1994). The limits of anchoring. *Journal of Behavioral Decision Making*, 7, 223–42.
- Chapman, G. B., & Johnson, E. J. (2002). Incorporating the irrelevant: anchors in judgments of belief and value. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.). *Heuristics and biases: The psychology of intuitive judgment*, 120–38.
- Doyon, M., Saulais, L., Ruffieux, B., & Bweli, D. (2015). Hypothetical Bias for Private Goods: Does Cheap Talk Make a Difference? *Theoretical Economics Letters*, 5(6), 749–56.
- Englich, B. & Mussweiler, T. (2001). Sentencing under Uncertainty: Anchoring Effects in the Courtroom. *Journal of Applied Social Psychology*, 31, 1535–51.

- Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: Differential effects of forewarning and incentives on self-generated and externally-provided anchors. *Journal of Behavioral Decision Making*, 18 (3), 199–212.
- Foster, H. & Burrows, J. (2017). Hypothetical bias: a new meta-analysis. In D. McFadden & K. Train (Eds.). *Contingent Valuation of Environmental Goods: A Comprehensive Critique*, 270–91.
- Fudenberg, D., Levine, D. K., & Maniadis, Z. (2012). On the Robustness of Anchoring Effects in WTP and WTA Experiments. *American Economic Journal: Microeconomics*, 4 (2), 131–45.
- Furnham, A. & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, 40 (1), 35–42.
- Gilbert, D. T., Pelham, B. W., & Krull, D. S. (1988). On cognitive busyness, when busy perceivers meet persons perceiving. *Journal of Personality and Social Psychology*, 54, 733–40.
- Green, D., Jacowitz, K. E., Kahneman, D., & McFadden, D. (1998). Referendum Contingent Valuation, Anchoring, and Willingness to Pay for Public Goods. *Resources and Energy Economics*, 20, 85–116.
- Ioannidis, K., Offerman, T., & Sloof, R. (2020). On the effect of anchoring on valuations when the anchor is transparently uninformative. *J Econ Sci Assoc*, 6, 77–94.
- Johnson, E. J., & Schkade, D. A. (1989). Bias in Utility Assessments: Further Evidence and Explanations. *Management Science*, 35 (4), 406–24.
- Jung, M. H., Perfecto, H., & Nelson, L. D. (2016). Anchoring in payment: Evaluating a judgmental heuristic in field experimental settings. *Journal of Marketing Research*, 53, 354–68.
- Kahneman, D., & Knetsch, J. (1993). Anchoring or Shallow Inferences: The Effect of Format. Unpublished manuscript. University of California, Berkeley.
- Koçaş, C., & Dogerlioglu-Demir, K. (2014). An empirical investigation of consumers' willingness-to-pay and the demand function: The cumulative effect of individual differences in anchored willingness-to-pay responses. *Marketing Letters* 25, 139–152.
- Kruger, J. (1999). Lake Wobegon be gone! The “below-average effect” and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology*, 77, 221–32.
- Li, L., Maniadis, Z., & Sedikides, C. (2020). Anchoring in Economics: A Meta-Analysis of Studies on Willingness-To-Pay and Willingness-To-Accept. *Journal of Behavioral and Experimental Economics*, 101629.
- List, J. A. (2001). Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards. *American Economic Review*, 91(5), 1498–507.
- List, J. A. (2003). Using Random nth Price Auctions to Value Non-Market Goods and Services. *Journal of Regulatory Economics*, 23, 193–205.

- List, J. A., & Gallet, C. A. (2001). What Experimental Protocols Influence Disparities Between Actual and Hypothetical Stated Values? *Environmental and Resource Economics*, 20(3), 241–54.
- Lorko, M., Servátka, M., & Zhang, L. (2018). Anchoring in Project Duration Estimation, MPRA Paper, 88456
- Maniadis, Z., Tufano, F., & List, J. A. (2014). One Swallow Doesn't Make a Summer: New Evidence on Anchoring Effects. *American Economic Review*, 104(1), 277–90.
- Mochon, D., & Frederick, S. (2013). Anchoring in sequential judgments. *Organizational Behavior and Human Decision Processes*, 122(1), 69–79.
- Moser, R., Raffaelli, R., & Notaro, S. (2014). Testing hypothetical bias with a real choice experiment using respondents' own money. *European Review of Agricultural Economics*, 41(1), 25–46.
- Murphy, J. J., Allen, P. G., Stevens, T. H., & Weatherhead, D. (2005). A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics*, (30) 3, 313–25.
- Murphy, J.J., & Stevens, T. H. (2004). Contingent Valuation, Hypothetical Bias, and Experimental Economics. *Agricultural and Resource Economics Review*, 33(2), 182–92.
- Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology*, 35, 136–64.
- Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin*, 26, 1142–50.
- Penn, J. M., & Hu, W. (2018). “Understanding hypothetical bias: An enhanced meta-analysis”. *American Journal of Agricultural Economics*, 100(4), 1186-1206.
- Plous, S. (1989). Thinking the unthinkable: The effect of anchoring on likelihood estimates of nuclear war. *Journal of Applied Social Psychology*, 19, 67–91.
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments & Computers*, 36(4), 717-731.
- Simmons, J. P., LeBoeuf, R. A., & Nelson, L. D. (2010). The effect of accuracy motivation on anchoring and adjustment: Do people adjust from provided anchors? *Journal of Personality and Social Psychology*, 99, 917-932.
- Simonson, I., & Drolet, A. (2004). Anchoring Effects on Consumers' Willingness-to-Pay and Willingness-to-Accept. *Journal of Consumer Research*, 31(3), 681–90.
- Strack, F., & Mussweiler, T. (1997). Explaining the Enigmatic Anchoring Effect: Mechanisms of Selective Accessibility. *Journal of Personality and Social Psychology*, 73(3), 437–46.
- Sugden, R., Zheng, J., & Zizzo, D. J. (2013). Not all anchors are created equal. *Journal of Economic Psychology*, 39(C), 21–31.

- Tufano, F. (2010). Are 'True' Preferences Revealed in Repeated Markets? An Experimental Demonstration of Context-Dependent Valuations. *Experimental Economics*, 13(1), 1–13.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–31.
- Wertenbroch, K., & Skiera, B. (2002). Measuring consumers' willingness to pay at the point of purchase. *Journal of Marketing Research*, 39, 228-241.
- Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, 125(4), 387–402.
- Yoon, S., Fong, N. M., & Dimoka, A. (2019). The robustness of anchoring effects on preferential judgments. *Judgment and Decision Making*, 14(4), 470.

Appendices

Appendix A: Experiment 1: transcript of instructions

Part 1: Introduction [all treatments, presented orally]

Welcome!

Thank you for taking part in our experiment! One randomly selected participant will receive a 50 PLN Rossmann voucher today.

Please turn off your mobile phones, remain silent, and do not communicate with one another. Conducting conversations, looking around the room, and showing your answers to others is not allowed.

If you have any questions during the experiment, raise your hand and wait for the experimenter to come to you. (Do not ask your question aloud!)

People who do not comply with these rules will be excluded from the experiment and lose the chance of winning the prize.

We would like to inform you that the study is anonymous and all data collected will be used solely for scientific purposes. Along with the next set of instructions, you will receive an individual respondent code that you will need until the end of the experiment. Place it in

a prominent spot on the bench/desktop in front of you and keep it until the end of the experiment.

Part 2: Rules [HypoLow, HypoHi]

In a moment, you will receive a questionnaire in which you will be asked to value a product. In the questionnaire, you will have to specify the maximum price you would be willing to pay for this product. Your reply will be purely declarative and no real transactions will be made on this basis. The amount you specify will not be disclosed to anyone.

If you have any questions, raise your hand and wait for the experimenter to approach you.

Part 2: Rules: [RealLow, RealHi]

In a moment, you will have the opportunity to buy a product. You will receive a valuation questionnaire in which you will be asked to specify the maximum price you would be willing to pay for this product. The amount you give will not be disclosed to anyone.

Your reply will be binding and may determine whether you are able to buy the product.

After everyone gives their amounts, three participants will be selected. If you are one of them, a transaction price will be drawn for you. If it is lower than or equal to the amount you specified, **you will be required to buy the product at the drawn price**; however, if the drawn transaction price is higher than the amount you provided, no transaction will take place.

The best thing you can do in this situation is to give your actual valuation, which is the maximum price you are willing to pay for the presented product. If you give an amount higher than your actual valuation, you may have to pay more than you are willing to. If, on the other hand, you give a lower valuation than your actual one, you may be disappointed with your inability to purchase the product at your actual price.

Example: The participant declares that the maximum price s/he is willing to pay for the product is 15 000 PLN; a price of 12 000 PLN is drawn. The participant buys the product for 12 000 PLN; however, if s/he declared only 11 000 PLN, s/he would not be able to buy the product. Of course, this is just an example; the amounts involved in the experiment will be significantly lower.

If you are required to make a purchase and you do not have enough money with you, you will be able to pay up to a week after the experiment; thus even if you do not have cash on you, you can still participate in the experiment.

If you have any questions, raise your hand and wait for the experimenter to approach you.

Part 3: Information about the product and valuation questionnaire [HypoLow, HypoHi]

Voucher for the professional caricature or portrait

The portrait or caricature will be made by a professional artist – Agnieszka Paczuska, a Warsaw portraitist and painter, with an arts education and 20 years' experience.

- a caricature or portrait showing you or a person of your choice (e.g., as a gift)
- A4 format
- technique: pencil/coal/mixed technique (black-and-white)
- drawn based on the photo you selected or in person in the artist's studio (in Warsaw) at a convenient time
- voucher to be valid for 3 months from the date of receipt

.....

Answer the following questions:

Would you, hypothetically, buy the presented voucher for 20 PLN [80 PLN] ?

- YES
- NO

Give the maximum price that, hypothetically, you would be willing to pay for the presented voucher. (Please enter a specific amount in PLN.)

Part 3: Information about the product and valuation questionnaire [RealLow, RealHi]

Voucher for the professional caricature or portrait

The portrait or caricature will be made by a professional artist – Agnieszka Paczuska, a Warsaw portraitist and painter, with an arts education and 20 years' experience.

- a caricature or portrait showing you or a person of your choice (e.g., as a gift)
- A4 format
- technique: pencil/coal/mixed technique (black-and-white)
- drawn based on a photo you have selected or in person in the artist's studio (in Warsaw) at a convenient time
- voucher to be valid for 3 months from the date of receipt

.....

Remembering that your answers are binding (and will depend on whether you are able to buy the product or not), answer the following questions:

Would you buy the presented voucher for 20 PLN [80 PLN] ?

- YES
- NO

Give the maximum price that you would be willing to pay for the presented voucher. (Please enter a specific amount in PLN.)

Appendix B: Experiment 1: transcript of the questionnaire

1. Have you ever bought a caricature or a portrait before?

- YES
- NO

2. What factor(s) did you take into consideration when evaluating the voucher during the experiment? (You can choose up to 3 relevant answers.)

- Work technique
- Work format
- Time to use the voucher
- Examples of the artist's works
- Attached information about the artist
- Whether you need this type of product
- How much you can afford to pay for it
- Other (please specify).....

3. Did you like the example works of the artist?

- Definitely yes
- Yes
- Neither yes nor no
- No
- Definitely not

4. How would you use the presented voucher?

- For myself
- As a gift
- I don't know

5. Are you interested in art?

- Definitely yes
- Yes
- Neither yes nor no
- No
- Definitely no

6. What do you think is the average market price of a black and white caricature in A4 format?

(enter a specific amount in PLN)

.....

7. What do you think is the average market price of a black and white portrait in A4 format?

(enter a specific amount in PLN)

.....

8. Gender

- Female
- Male

9. Year of birth

.....

10. Are you in a relationship?

- Yes, I have a boyfriend/girlfriend
- Yes, I have a fiancé/fiancée
- Yes, I have a husband/wife
- No, I am single

11. Do you work currently?

- Yes
- No

12. How would you evaluate your financial situation?

- I don't have enough money for the most urgent needs
- I have to deny myself many things to keep enough money for living
- I have enough money for everyday living, but I can't afford higher expenses
- I have enough money for all expenses, and I can save
- I am wealthy, I do not have to save even for larger expenses
- I don't know/it's hard to say

13. Where were you born?

- a village
- a city with less than 100,000 residents
- a city with over 100,000 residents

Appendix C. Experiment 1: variable labels

hypothetical: 1 - for declarative (hypothetical) valuation, 0 - for real valuation

high_anchor: 1 - for a high anchor, 0 - for a low anchor

hypo#high_a: interaction between *hypothetical* and *high_anchor*

male: 1 - male, 0 - female

age

relationship: 1 - if the participant is in a relationship, 0 - if the participant is single

unemployed: 1 - if the participant is unemployed, 0 - otherwise

city: 1 - if the participant's place of birth is a city, 0 - if the participant's place of birth is a village

fin_situation : 1 - if the participant is in a good or very good financial situation, 0 - in all other cases

technique_e : 1 - if the participant took the work technique into consideration in the valuation process, 0 - otherwise

format_e : 1- if the participant took the work format into consideration in the valuation process, 0 - otherwise

time_e 1- if the participant took the time to use the voucher into consideration in the valuation process, 0 - otherwise

examples_e 1- if the participant took the examples of the artist's works into consideration in the valuation process, 0 - otherwise

artist_e: 1- if the participant took the information about the artist into consideration in the valuation process, 0 - otherwise

needs_e: 1- if the participant took his/her needs into consideration in the valuation process, 0 - otherwise

finance_e: 1- if the participant took into consideration how much (s)he can afford to pay in the valuation process, 0 - otherwise

others_e: 1- if the participant took other factors into consideration in the valuation process, 0 - otherwise

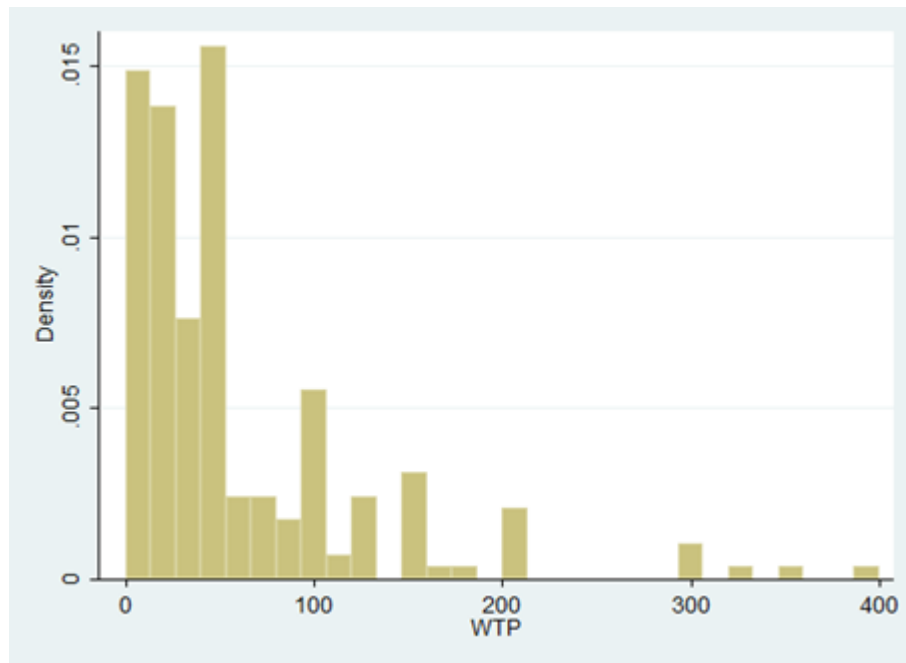
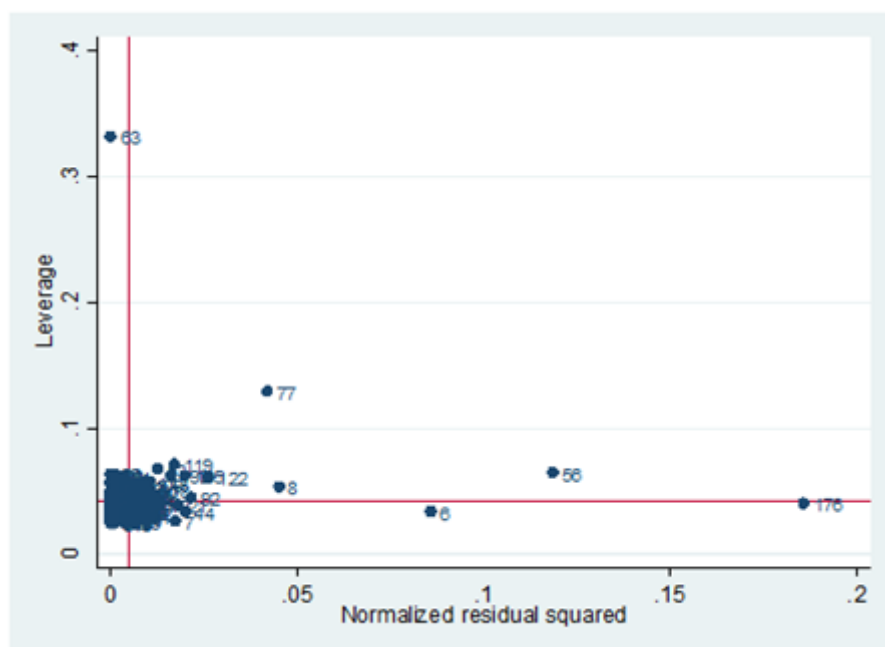
examples: 1- if the participant likes or really likes the examples of the artist's work, 0 - otherwise

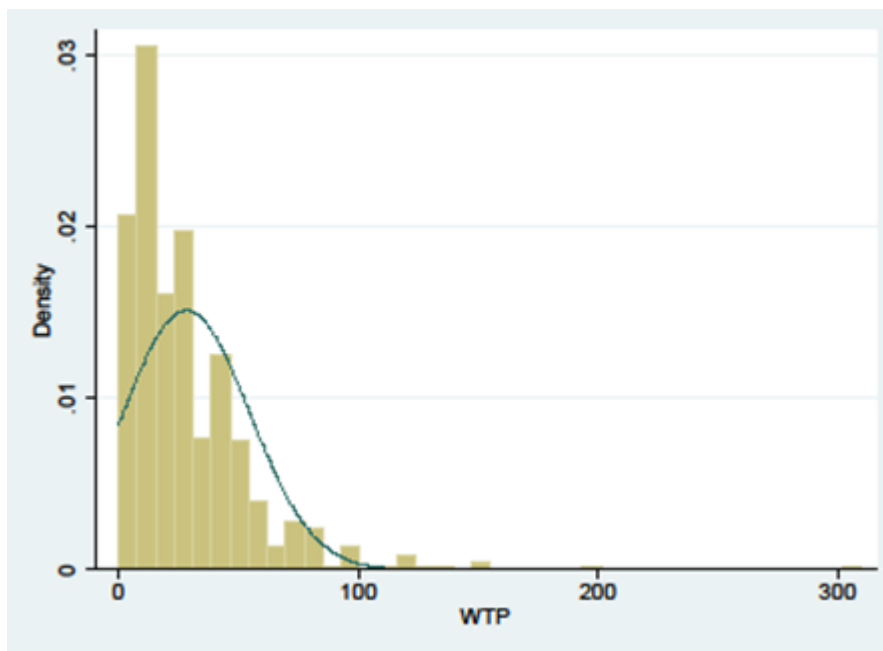
gift: 1 - if the participant would like to use the voucher as a gift, 0 - in all other cases

art: 1- if the participant is interested in art, 0 - otherwise

price_caricature: the perceived market price of a caricature

price_portrait: the perceived market price of a portrait

Appendix D. Experiment 1: figures**Fig. 4** Experiment 1: histogram of WTP values across the sample**Fig. 5** Experiment 1: The Leverage and Squared Normalized Residuals

Appendix E. Experiment 2: figures**Fig. 6** Experiment 2: selected products.**Fig. 7** Experiment 2: histogram of WTP values across the sample

Appendix F: Experiments 1 and 2: transcript of instructions

Part 1: Introduction [all treatments, presented orally]

Welcome!

We are researchers from the University of Warsaw. We are conducting a survey, which will only take a few minutes. I am wondering if you would like to participate. At the end of this survey, you will play a game in which you can win a Westfield Arkadia Gift Card, worth 100 PLN. If you have any questions during the survey, do not hesitate to ask.

We would like to inform you that the study is anonymous and all data collected will be used solely for scientific purposes.

Part 2: Rules [HypoLow, HypoHi]

In a moment, you will receive a questionnaire in which you will be asked to value the presented product. In the questionnaire, you will have to specify the maximum price you would be willing to pay for this product. It will be purely declarative and no real transactions will be made on this basis.

Part 2: Rules: [RealLow, RealHi]

In a moment, you will have the opportunity to buy the presented product. You will receive a valuation questionnaire in which you will be asked to specify the maximum price you would be willing to pay for this product.

Your choice will be binding: it will determine whether you buy the product. After you state your amount, a transaction price will be drawn for you. If it is lower than or equal to the amount you specified, **you will be required to buy the product at the drawn price**; however, if the drawn transaction price is higher than the amount you provided, no transaction will take place.

The best thing you can do in this situation is to give your actual valuation, which is the maximum price you are willing to pay for the presented product. If you give an amount higher than your actual valuation, you may have to pay more than you are willing to. If, on the other hand, you give a lower valuation than your actual one, you may be disappointed with your inability to purchase the product offered at an acceptable price.

Example: The participant declares that the maximum price s/he is willing to pay for the product is 15 000 PLN; a price of 10 000 PLN is drawn. The participant buys the product for 10 000 PLN; however, if the price was 20 000 PLN, s/he would not be able to buy the product. Of course, this is just an example; the amounts involved in the experiment will be significantly lower.

Part 3: Information about the product and valuation questionnaire [HypoLow, HypoHi]

[Short information about the product]

Answer the following questions:

Would you, hypothetically, buy the selected product for 10 PLN [60 PLN]?

- YES
- NO

Give the maximum price that, hypothetically, you would be willing to pay for the presented product. (Please enter a specific amount in PLN.)

Part 3: Information about the product and valuation questionnaire [RealLow, RealHi]

[Short information about product]

Remembering that your answers are binding and whether you buy the product will depend on them, answer the following questions:

Would you buy the presented product for 10 PLN [60 PLN]?

- YES
- NO

Give the maximum price that you are willing to pay for the presented product. (Please enter a specific amount in PLN.)

Appendix G: Experiments 1 and 2: transcript of a questionnaire

1. How would you use the presented product?

- For myself
- As a gift
- I don't know

2. Do you like this product?

- Definitely yes
- Yes
- Neither yes nor no
- No
- Definitely not

3. Gender

- Female
- Male

4. Year of birth

.....

5. Education level:

- Primary

- Vocational
- Secondary
- Higher

6. Do you work currently?

- Yes
- No

Appendix H. Experiments 2 and 3: variable labels

hypothetical: 1 - for declarative (hypothetical) valuation, 0 - for real valuation

high_anchor: 1 - for a high anchor, 0 - for a low anchor

hypo#high_a: interaction between *hypothetical* and *high_anchor*

male: 1 - male, 0 - female

age

education: 1 - if the participant has higher or secondary education, 0 - if the participant has primary or vocational education

higher_edu: 1 - if the participant has higher education, 0 - in all other cases

unemployed: 1 - if the participant is unemployed, 0 - otherwise

gift: 1 - if the participant would like to use the mug/honey as a gift, 0 - in all other cases

very_nice: 1 - if the participant assesses the presented mugs as very nice, 0 - in all other cases

attractive: 1 - if the selected honey is attractive or very attractive for the participant, 0 - in all other cases

likes_very_much: 1 - if the participant likes honey very much, 0 - in all other cases

experimenter: 1 - if the experiment was conducted by experimenter no. 1, 0 - if the experiment was conducted by experimenter no. 2

mug_elephant: 1 - if participant selected the mug with an elephant; 0 - otherwise

mug_cat: 1- if participant selected the mug with a cat, 0 - otherwise

honey_ginger: 1- if participant selected the honey with garlic and ginger; 0 - otherwise

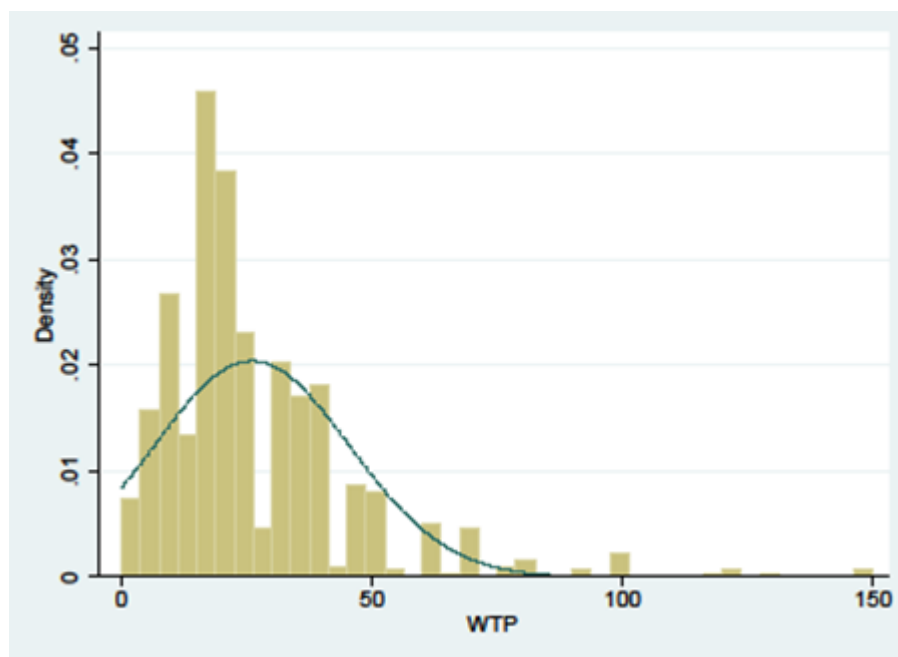
honey_cocoa: 1- if participant selected the honey with cocoa, 0 - otherwise

Appendix I. Experiment 3: figures

Fig. 8 Experiment 3: selected products



Fig. 9 Experiment 3: histogram of the WTP values across the sample





UNIVERSITY OF WARSAW
FACULTY OF ECONOMIC SCIENCES
44/50 DŁUGA ST.
00-241 WARSAW
WWW.WNE.UW.EDU.PL